Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Autotracing of *Escherichia coli* acetate CoA-transferase α-subunit structure using 3.4 Å MAD and 1.9 Å native data

S. Korolev, O. Koroleva, K. Petterson, M. Gu, F. Collart, I. Dementieva and A. Joachimiak*

Biosciences Division and Structural Biology Center, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, IL 60439, USA

Correspondence e-mail: andrzejj@anl.gov

Received 23 March 2002
Accepted 19 September 2002

PDB Reference: acetate CoA-transferase α-subunit, 1k6d, r1k6dsf.

The automation of protein structure determination is an essential component for high-throughput structural analysis in protein X-ray crystallography and is a key element in structural genomics. This highly challenging undertaking relies at present on the availability of high-quality native and derivatized protein crystals diffracting to high or moderate resolution, respectively. Obtaining such crystals often requires significant effort. The present study demonstrates that phases obtained at low resolution (>3.0 Å) from crystals of SeMet-labeled protein can be successfully used for automated structure determination. The crystal structure of acetate CoA-transferase α-subunit was solved using 3.4 Å multi-wavelength anomalous dispersion data collected from a crystal containing SeMet-substituted protein and 1.9 Å data collected from a native protein crystal.

1. Introduction

High-throughput protein crystallography is feasible because of dramatic progress in the automation of gene cloning, protein purification, crystallization and particularly in utilization of multiwavelength anomalous dispersion (MAD) methods using synchrotron radiation (Hendrickson, 2000; Hendrickson *et al.*, 1985) and the development of new software for automated protein-structure solution (Lamzin & Perrakis, 2000; Perrakis *et al.*, 1999; Terwilliger & Berendzen, 1999). TProjects like the NIH-sponsored Protein Structure Initiative (Norvell & Machalek, 2000; Stevens *et al.*, 2001) foster the development of instrumentation, methods and software to automate the process of protein-structure determination. However, a significant bottleneck is the interpretation of electron-density maps and model building. Current autotracing programs can be used only with high-resolution diffraction data and with moderate resolution phases and are therefore strongly dependent on the availability of high-quality crystals and synchrotron facilities.

Since crystallization screens rarely produce diffraction-quality protein crystals, significant effort is needed to refine crystallization conditions in order to produce crystals suitable for structure determination. The need to prepare crystals of derivatized protein further increases the required effort. Additional difficulties arise from the needs of the MAD experiment, which requires larger amounts of data than experiments with native protein crystals and is more vulnerable to radiation damage. Therefore, methods that enable automated structure determination using lower resolution MAD data would be of general interest to the structural

biology community and will expand the spectrum of targets suitable for high-throughput crystallography.

In this paper, we present the crystal structure of the *Escherichia coli* acetate CoA-transferase (ACT) α-subunit solved with minimal manual intervention using only low-resolution (3.4 Å) MAD data collected from the crystal of SeMet-derivatized protein and 1.9 Å resolution data obtained from the crystal of native protein. The enzyme belongs to the family of coenzyme A-transferases found in archaea, bacteria, metazoa and mitochondria of mammals. The structure is of biological relevance, as *E. coli* ACT shows 36% sequence identity with the product of the human gene encoding mitochondrial succinyl-CoA:3-oxoacid CoA-transferase (SCOT; Jencks, 1973). Mutations of the latter enzyme cause a metabolic disease which leads to severe ketoacidosis.

2. Materials and methods

2.1. Sample preparation

All reagents were purchased from Sigma, except where specifically mentioned otherwise in the text. The open reading frame of the *E. coli* ACT α-subunit was amplified from genomic DNA with high-fidelity *Pfx* DNA polymerase, using conditions and reagents provided by the vendor (Invitrogen, Carlsbad, CA, USA). The gene was cloned into pET30XaLIC (Novagen, Madison, WI, USA) using the Ligation Independent Cloning protocol (Dieckman *et al.*, 2002). This process generated an expression clone producing a fusion protein with a thrombin-cleavable His tag and a factor Xa-cleavable S tag (His-S-ACT). The fusion protein was overexpressed in *E. coli* BL21-Gold (DE3) (Stratagene) harboring an extra plasmid encoding three rare tRNAs (AGG and AGA for Arg; ATA for Ile). The cells were grown in LB media at 310 K to an OD_{600} of $\sim$0.6 and following induction with 1 m*M* IPTG for 4 h at 303 K the bacteria were harvested and stored at 203 K. The frozen cells were thawed, resuspended in binding buffer (300 m*M* NaCl, 5% glycerol, 50 m*M* sodium phosphate pH 8.0 and 10 m*M* imidazole) and lysed by sonication after the addition of lysozyme to 1 mg ml^{-1} and 1 m*M* PMSF. The lysate was clarified by centrifugation, passed through a 0.2 µm filter and applied to a metal-chelate Ni–NTA affinity column (Qiagen). After the column was washed, His-S-ACT was eluted from the column with 250 m*M* imidazole in binding buffer and dialyzed into thrombin cleavage/capture buffer *A* (20 m*M* Tris–HCl pH 8.4, 0.15 m*M* NaCl, 2.5 m*M* $CaCl_2$). The His tag was cleaved from the protein by treatment with biotinylated thrombin (Novagen) overnight at room temperature using 1 unit of protease per milligram of protein. The mixture was incubated with Streptavidin Agarose (Novagen) at room temperature for 30 min and the biotinylated thrombin was removed from the cleaved protein by centrifugation at 500*g* for 5 min in a Spin Filter. The cleaved protein (S-ACT) was then purified from the His-S-ACT by passage of the mixture through a second Ni–NTA column. The S tag was cleaved by treatment with factor Xa protease (Novagen) in buffer *A* overnight at room temperature using 2 units of protease per 50 µg of protein. The mixture was then incubated with Xarrest Agarose (Novagen) at room temperature for 30 min and the cleaved protein was purified from factor Xa protease by centrifugation at 500*g* for 5 min using a Spin Filter. ACT was further purified using a Mono Q column (Pharmacia) equilibrated with 20 m*M* bis-tris propane pH 8.0, 50 m*M* NaCl, 2 m*M* DDT buffer and eluted with a 0.05–0.5 *M* NaCl gradient. The protein was dialyzed against 20 m*M* bis-tris propane pH 8.0, 100 m*M* NaCl, 2 m*M* DDT and concentrated using a BioMax concentrator (Millipore). SeMet-labeled protein was prepared as in Walsh *et al.* (1999).

Table 1
Data collection and refinement statistics.

Values for highest resolution shell (1.9–1.97 Å for native crystals and 3.4–3.52 Å for SeMet-containing crystals) are shown in parentheses.

	Wavelength (Å)	Resolution (Å)	R_{merge} (%)	Completeness (%)	$I/\sigma(I)$ at high resolution
Crystal 1 (SeMet)					
Peak	0.97936	50–3.4	9.7 (43.1)	99.9 (99.9)	4.7
Inflection	0.97950	50–3.4	11.1 (56.9)	99.9 (99.9)	3.1
Low remote	1.00800	50–3.4	10.8 (58.0)	99.7 (99.0)	1.9
Crystal 2 (native)	0.97934	50–1.9	7.8 (67.7)	100 (100)	3

Refinement.

Resolution (Å)	50–1.9
No. of protein non-H atoms	3238
No. of Mg atoms	3
No. of water molecules	228
No. of reflections	36763
No. of test-set reflections (5%)	1823
R (%)	20.1 (25.4)
Free R (%)	23.5 (29.7)
R.m.s.d. bonds (Å)	0.009
R.m.s.d. angles (°)	1.5
Overall B factor (Å^2)	30.8
Ramachandran plot†	
Most favored regions (%)	93.9
Allowed regions (%)	4.4
Generously allowed regions (%)	1.7
Disallowed regions (%)	0.0

† Ramachandran plot parameters were calculated using *PROCHECK* (Laskowski *et al.*, 1993).

2.2. Crystallization

Hexagonal crystals of S-ACT protein were grown at room temperature using the hanging-drop vapor-diffusion method in 6–10% MPD, 0.1 *M* sodium citrate pH 6.5, 0.2 *M* magnesium acetate. Crystals reached maximum dimensions of 0.3 × 0.1 × 0.1 mm in one week. For data collection, crystals were gradually transferred into a solution containing 20% MPD and flash-frozen in liquid nitrogen. The crystals belonged to space group $P6_2$, with unit-cell parameters $a = b = 88.9$, $c = 106.6$ Å and two molecules per asymmetric unit, and diffracted to 3.2 Å resolution. Crystals of SeMet-labeled S-ACT protein (0.15 × 0.07 × 0.07 mm) were grown under similar conditions. MAD data were collected to 3.4 Å resolution at three wavelengths near the selenium absorption edge (Table 1).

121 Small crystals of native ACT with both His and S tags cleaved off were grown under slightly altered conditions (6% MPD, 0.1 *M* glycine pH 8.0 and 0.2 *M* magnesium acetate) and reached maximum size (0.1 × 0.03 × 0.03 mm) in three months. In spite of their small size, they diffracted to 1.9 Å resolution and had unit-cell parameters $a = b = 88.5$, $c = 105.1$ Å. All data were collected at beamline 19ID of the Structural Biology Center at the Advanced Photon Source, Argonne National Laboratory and were processed with *HKL*2000 (Otwinowski & Minor, 1997). Details of data collection and processing are shown in Table 1.

2.3. Structure determination

Selenium positions were found with the program *Shake-and-Bake* (Weeks & Miller, 1999) using the data set collected at the absorption-peak energy. Phases were calculated with the program *SHARP* (de La Fortelle & Bricogne, 1997) following density modification with the program *SOLOMON* (Abrahams & Leslie, 1996) as implemented in *SHARP*. Se-atom positions were subjected to the program *FINDNCS* (Lu, 1999) from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994) to find the non-crystallographic symmetry (NCS) matrix and to automatically build the mask with the program *NCSMASK*. The resulting phases were combined with the high-resolution native data set and phases were extended to a maximum resolution of 1.9 Å with program *DM* (Cowtan & Zhang, 1999) using NCS averaging. The final phases were used in the program *ARP*/*wARP* (Perrakis *et al.*, 1999) for automated building of the protein structure.

Atomic model and electron-density maps were visualized using the program *O* (Jones, 1985) and refinement was carried out with the program *CNS* (Brünger *et al.*, 1998).

The correlation coefficient between different electron-density maps and final protein model was calculated using *CCP*4 programs as implemented in the 'Map correlation' task in the ccp4i interface v. 1.1.1.

3. Results and discussion

3.1. Crystallization and data collection

32 mg of ~95% pure His-S-ACT protein was obtained from 1 l of culture after a single immobilized metal-affinity chromatography (IMAC) step. A portion of the His-S-ACT protein was used for crystallization screens, while the remainder was subjected to thrombin cleavage. After cleavage of the His tag and further purification using IMAC, the S-ACT protein was more than 98% pure and was also subjected to crystallization screening. The efficiency of S-tag cleavage with factor Xa was low and we did not use the native ACT protein for crystallization screening. Both His-S-ACT and S-ACT

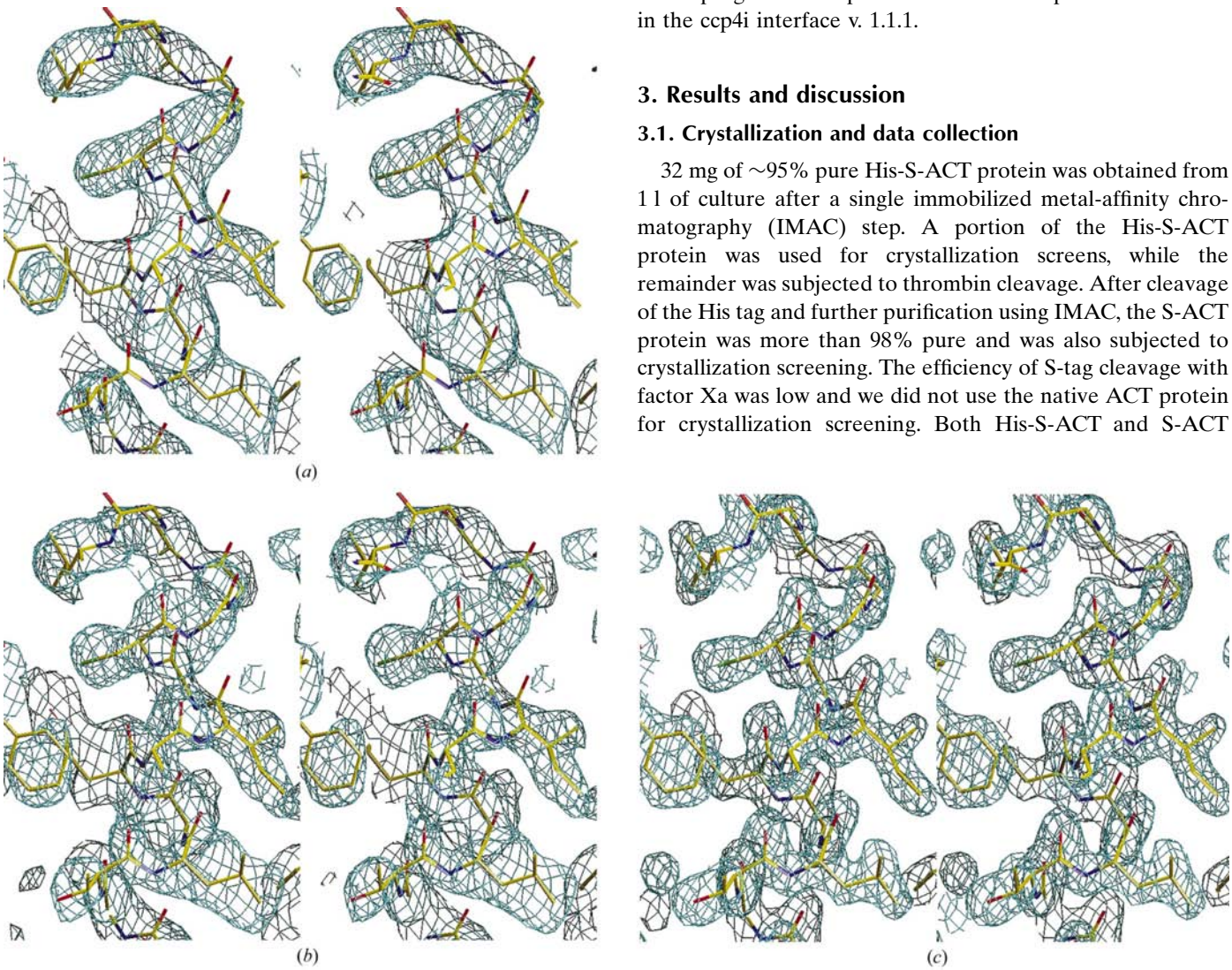

Figure 1
Stereoview of electron-density maps obtained at different stages of structure solution. A representative part of the structure is shown (*a*) with electron density obtained with *SHARP*/*SOLOMON* at 3.4 Å resolution, (*b*) with electron density obtained after phase extension to 1.9 Å and averaging with the program *DM* and (*c*) with $2F_o - F_c$ electron density.

Table 2
Phase calculation, modification and autotracing statistics.

SHARP		*DM*				*ARP/wARP*	
Resolution (Å)	No. of wavelengths	Resolution (Å)	NCS	No. of cycles	Correlation coefficient	No. of cycles to obtain 0.9 connectivity	No. of amino acids traced in 500 cycles
3.4	2	1.9	+	100	0.71	70	400
3.4	2	1.9	+	10	0.64	150	400
3.4	2	2.1	+	100	0.68	180	370
4.0	2	1.9	+	100	0.56	670	390†
3.4	1	1.9	+	100	0.65	90	400
3.4	2	2.3	+	100	0.65	(0.64)‡	160
3.4	3	1.9	+	100	0.32	(0.54)‡	70
3.4	2	1.9	−	8§	0.51	(0.52)‡	60
3.4	2	1.9	−	100	0.49	(0.50)‡	50

† Traced in 1000 cycles. ‡ Maximum connectivity achieved in 500 cycles is shown in parentheses. § *DM* protocol with AUTO mode for number of cycles.

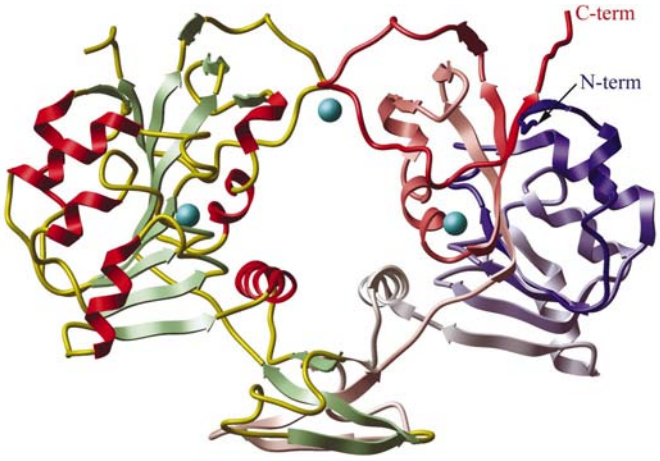

Figure 2
Ribbon representation of the ACT α-subunit dimer. One monomer is color coded according to the secondary structure, with β-strands in green, α-helices in red and loop regions in yellow. The second monomer is gradient color coded from blue at the N-terminus to red at the C-terminus. Three Mg^{2+} ions are shown as cyan spheres. Two of the Mg^{2+} ions are bound to identical sites in molecules *A* and *B* and one is bound at the monomer interface near the non-crystallographic symmetry twofold axes. The figure was produced with the program *ICM* (Abagyan & Totrov, 2000).

proteins produced crystals under similar conditions and diffracted only to ~3.2 Å. Although larger crystals were obtained after refinement of crystallization conditions, the diffraction limit did not improve.

SeMet-labeled His-S-ACT protein was prepared after the first crystals of native protein were obtained. The yield of purified SeMet-labeled His-S-ACT protein was much lower than that of native protein and only 2.2 mg of pure S-ACT protein was obtained after thrombin cleavage and a second IMAC purification. The SeMet S-ACT derivative was crystallized under similar conditions to the native protein. A 3.4 Å resolution MAD data set was collected at three different wavelengths from a small crystal (0.15 × 0.07 × 0.07 mm) of SeMet-labeled S-ACT (Table 1). The data statistics were considerably poorer for the third data set collected at the low-energy remote. The correlation coefficients between data sets, calculated with *SOLVE*, were 0.77, 0.07 and 0.04 for the first *versus* second, first *versus* third and second *versus* third data sets, respectively.

In a parallel effort, the native ACT protein was produced as described in §2 and was further purified on a Mono Q column, yielding a sample virtually free of contaminants. Small crystals (0.1 × 0.03 × 0.03 mm) of this protein appeared in three months and a complete data set was collected to 1.9 Å resolution from a single crystal.

3.1.1. Phasing and autotracing. The 1.9 Å data obtained from the native protein crystal and the 3.4 Å MAD data enabled determination of the ACT α-subunit structure. The important question for high-throughput crystallography was whether the data were suitable for automated structure determination. The crystal contains two protein molecules per asymmetric unit, with six methionines per molecule. Ten heavy-atom peaks were found using the program *Shake-and-Bake* with a data set collected at the selenium white line. Initial phases were calculated with the program *SHARP* followed by density modification with *SOLOMON*. The low-resolution maps obtained by this process were of good quality and clearly showed recognizable α-helices and β-strands of the protein structure (Fig. 1*a*). Phase extension to the highest resolution of the native data (1.9 Å) produced maps with more features (Fig. 1*b*). Nevertheless, the quality of extended phases was insufficient for autotracing with the program *ARP/wARP* (Table 2) and we utilized the presence of non-crystallographic symmetry (NCS). Heavy-atom positions were used with *FINDNCS* to produce an NCS matrix. The mask was built automatically by surrounding heavy-atom positions of one group with 20 Å spheres and was refined with the program *NCSMASK*.

Phase extension was performed with the program *DM* using NCS averaging. Final phases were of sufficient quality for automated determination of the structure using *ARP/wARP* (Table 2). 30 missing residues in one of the monomers were built simply by copying the corresponding residues from an NCS-related molecule built with *ARP/wARP*. Thus, all steps of structure determination from phasing to a 98% complete model were carried out with minimal manual intervention.

To evaluate the limitations of this automated structure-determination approach, we altered the parameters in *SHARP* and *DM* as shown in Table 2. In *SHARP* we used one, two or three data sets collected at different wavelengths and applied 3.4 or 4.0 Å resolution cutoffs. In *DM* we used the default protocol with solvent flattening, histogram mapping, multi-resolution modification and resolution-dependent perturbation gamma correction, and varied the number of cycles, the high-resolution limit and the NCS averaging. All parameters in *ARP/wARP* were set to the default values suggested in the *ARP/wARP* manual (http://www.embl-hamburg.de/ARP/manual) and the high-resolution

limits for the native data were the same as shown in Table 2 for *DM*.

The upper part of Table 2 summarizes the statistics for successful runs, while the bottom three rows correspond to situations where autotracing failed. The calculations with 2.3 Å resolution data were partially successful. Radiation damage to the crystal was clearly observed during processing of the diffraction data. Phasing with three-wavelength data sets failed.

The best results were obtained using two wavelengths (peak and inflection) at 3.4 Å resolution for phasing and phase extension to the highest resolution with 100 cycles of density modification and NCS averaging. The resulted phases were of sufficient quality for quick convergence of autotracing. The single anomalous dispersion (SAD; Wang, 1985) approach in which only data collected at peak energy were used for phasing also produced phases of a good quality sufficient for quick autotracing. Using only 2.1 Å data for phase extension and autotracing as well as using 4 Å resolution for phase calculation with subsequent phase extension to the highest resolution also allowed successful autotracing of the model. Autotracing failed in all cases when NCS averaging was omitted during phase extension.

3.1.2. Biological significance of the structure. *E. coli* acetate CoA-transferase shares a high degree of sequence similarity with other members of the coenzyme A-transferase family, which consists of 71 sequence homologs in 50 unique species of bacteria, archaea and metazoa. The enzyme catalyzes the transfer of CoA from an acyl-CoA donor to a carboxylate acceptor and plays a key role in prokaryotic and eukaryotic metabolism, enabling the activation of various carbon compounds for subsequent oxidation (Jenkins & Weitzman, 1986). The bacterial enzymes are heterotetrameric ($\alpha_2\beta_2$), in which the α-subunit binds CoA while the β-subunit provides catalytic glutamate. The ACT α-subunit shows 36% sequence identity with the N-terminal half of the mitochondrial enzyme succinyl-CoA:3-ketoacid CoA transferase (SCOT; EC 2.8.3.5), which consists of a single chain collinear with the two α/β bacterial subunits and is essential in ketone metabolism (Rochet & Bridger, 1994).

The three-dimensional structure of the ACT α-subunit (Fig. 2) and its dimeric assembly show overall structural similarity to the α-subunits of glutaconate CoA-transferase (GST) from *A. frementans* (Jacob *et al.*, 1997; PDB code 1poi), the only CoA-transferase with a known three-dimensional structure. The r.m.s. deviation between ACT and GST α-subunits is 1.9 Å for most C^{α} atoms of ACT and the orientation of the subunits in a dimer is very similar. Loops 77–82, 132–151 and the C-terminal residues 245–319 of GST do not have structural analogs in ACT. Of the four GST residues predicted to be involved in substrate binding, only Gln103 is preserved, while Thr27 is substituted with Met, Tyr74 with His and Ser78 belongs to the loop which is missing in ACT. These changes are likely to reflect the different substrate specificity of the enzymes. Three Mg^{2+} ions were found in the ACT structure (Fig. 2). None of them are located near the putative active site, hence their functional relevance is not clear. The CoA-binding signature (pfam01144, Pfam v6.6; Sonnhammer *et al.*, 1997) common to ACT and SCOT varies in GST. The GST α-subunit shows less than 20% sequence identity with SCOT, while the ACT sequence has 36% identity with SCOT. This suggests the ACT structure may be a useful intermediate in the modeling of the mitochondrial enzyme. Neither ACT nor GST was crystallized with substrates and additional structural and modeling experiments are required to further study the catalytic mechanism.

4. Conclusions

Structural genomics strategies are often based on targeting the largest possible number of proteins in order to increase the overall success of the project. Current technologies enable a high-throughput approach for the initial steps from cloning to initial crystallization screening for the majority of soluble proteins. The later steps of structure determination have considerably lower productivity and in general only a limited set of proteins suitable for high-throughput structure determination are selected to continue work with. At this point, it is important to know the current limitations of automated structure-determination methods. Here, we have showed that density modification and extension of low-resolution phases, which are routinely applied in conventional protein crystallography (one of the best examples is in Braig *et al.*, 1994), can be successfully applied to high-throughput automated structure determination. Phases calculated from 3.4 Å MAD data and extended to 1.9 Å resolution for the native data were successfully used for the automated structure determination of a medium-sized protein. We also obtained an automated solution in test cases when initial phases were only calculated at 4.0 Å resolution or with phase extension to 2.1 and 2.3 Å. These data cannot truly simulate situations in real protein crystals with such diffraction limits, but they also provide evidence that 3.4 Å MAD and 1.9 Å native data are not at the threshold level of automated structure determination.

An essential element of our approach was NCS averaging during phase extension. Without NCS averaging, the data were not sufficient to enable autotracing with *ARP*/*wARP*. Similarly to previous observations (Rice *et al.*, 2000), we found that the single-wavelength anomalous dispersion approach (SAD) also provides phases of a good quality and enabled quick autotracing of the structure. In addition to broadening the spectrum of protein crystals suitable for high-throughput structural analysis, these results are also important for choosing a data-collection strategy. Since without NCS averaging resolution limits became a crucial parameter, the best strategy seems to be collecting SAD data of the highest possible resolution at the absorption-peak energy. On the other hand, MAD phases are generally of a better quality than SAD phases (Rice *et al.*, 2000). In this respect, an alternative strategy could require the collection of data at two different wavelengths: first a quick data set at lower than the absorption-edge energy (for example, at the inflection-point energy) to minimize radiation damage and second a data set at peak

energy with the highest possible resolution. The first data set will serve as a test for optimization of the data-collection strategy and will provide data for resolving the initial phase ambiguity of SAD phases. In order to evaluate these and other strategies, more crystals of different quality have to be tested. Structural genomics projects should provide an excellent opportunity for such experiments.

We wish to thank all members of the MCSG group in the Biosciences Division, the Structural Biology Center at Argonne National Laboratory and Lindy Keller for help in preparation of this manuscript. This work was supported by the National Institutes of Health Grant GM62414 and the US Department of Energy, Office of Biological and Environmental Research under Contract W-31–109-ENG-38.

References

Abagyan, R. & Totrov, M. (2000). *ICM 2.8 User's Reference.* MolSoft, La Jolla, USA.
Abrahams, J. P. & Leslie, A. G. W. (1996). *Acta Cryst.* D**52**, 30–42.
Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D. C., Joachimiak, A., Horwich, A. L. & Sigler, P. B. (1994). *Nature (London)*, **371**, 578–586.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cowtan, K. D. & Zhang, K. Y. (1999). *Prog. Biophys. Mol. Biol.* **72**, 245–270.
Dieckman, L., Gu, M., Stols, L., Donnelley, M. I. & Collart, F. R. (2002). *Protein Expr. Purif.* **25**, 1–7.
Hendrickson, W. A. (2000). *Trends Biochem. Sci.* **25**, 637–643.
Hendrickson, W. A., Smith, J. L. & Sheriff, S. (1985). *Methods Enzymol.* **115**, 41–55.
Jacob, U., Mack, M., Clausen, T., Huber, R., Buckel, W. & Messerschmidt, A. (1997). *Structure*, **5**, 415–426.
Jencks, W. P. (1973). *The Enzymes*, edited by P. D. Boyer, pp. 483–496. New York: Academic Press.
Jenkins, T. M. & Weitzman, P. D. (1986). *FEBS Lett.* **205**, 215–218.
Jones, T. A. (1985). *Methods Enzymol.* **115**, 157–171.
de La Fortelle, E. & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.
Lamzin, V. S. & Perrakis, A. (2000). *Nature Struct. Biol.* **7** (*Suppl.*), 978–981.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lu, G. (1999). *J. Appl. Cryst.* **32**, 365–368.
Norvell, J. C. & Machalek, A. Z. (2000). *Nature Struct. Biol.* **7** (*Suppl.*), 931.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.
Rice, L. M., Earnest, T. N. & Brunger, A. T. (2000). *Acta Cryst.* D**56**, 1413–1420.
Rochet, J. C. & Bridger, W. A. (1994). *Protein Sci.* **3**, 975–981.
Sonnhammer, E. L., Eddy, S. R. & Durbin, R. (1997). *Proteins*, **28**, 405–420.
Stevens, R. C., Yokoyama, S. & Wilson, I. A. (2001). *Science*, **294**, 89–92.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.
Walsh, M. A., Dementieva, I., Evans, G., Sanishvili, R. & Joachimiak, A. (1999). *Acta Cryst.* D**55**, 1168–1173.
Wang, B.-C. (1985). *Methods Enzymol.* **115**, 90–111.
Weeks, C. M. & Miller, R. (1999). *Acta Cryst.* D**55**, 492–500.