Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Generating isomorphous heavy-atom derivatives by a quick-soak method. Part II: phasing of new structures

Peter D. Sun* and Sergei Radaev

Structural Immunology Section, Laboratory of Immunogenetics, National Institute of Allergy and Infectious Diseases, National Institutes of Health, 12441 Parklawn Drive, Rockville, MD 20852, USA

Correspondence e-mail: psun@nih.gov

Received 1 January 2002
Accepted 10 April 2002

A quick-soak method has been applied to generate *de novo* heavy-atom phasing to solve two new protein structures, a type II transforming growth factor β receptor (TBRII) and a natural killer cell receptor–ligand complex, NKG2D–ULBP3. In the case of TBRII, a crystal derivatized for only 10 min in saturated $HgCl_2$ provided adequate phasing for structure determination. Comparison between $HgCl_2$ derivatives generated by 10 min soaking and by 12 h soaking revealed similar phasing statistics. The shorter soak, however, resulted in a derivative more isomorphous to the native than the longer soak as judged by changes in the unit-cell parameter *a* upon derivatization as well as by the quality of a combined SIRAS electron-density map. In the case of the NKG2D–ULBP3 structure, all overnight soaks in heavy-atom solutions resulted in crystal lattice disorder and only the quick soaks preserved diffraction. Despite fragile lattice packing, the quick-soaked K_2PtCl_4 derivative was isomorphous with the native crystal and the electron-density map calculated from combined SIR and MAD phases is better than that calculated from MAD phases alone. Combined with mass-spectrometry-assisted solution heavy-atom derivative screening and the use of synchrotron radiation, the quick-soak derivatization has the potential to transform the time-consuming conventional heavy-atom search into a real-time 'on-the-fly' derivatization process that will benefit high-throughput structural genomics.

1. Introduction

We have tested the ability to generate heavy-metal derivatives by a quick-soak method on hen egg-white lysozyme and Fc receptor crystals (Sun *et al.*, 2002). The success of the method relies on the rapid rate of derivatization reactions, most of which occur in minutes to hours rather than days. For non-covalent derivatizations, in which the rate-limiting step is that of heavy-atom diffusion in crystals, a complete derivatization can occur in less than 1 min (Nagem *et al.*, 2001). For covalent adducts, both the diffusion rate and the rate of chemical reactions affect the extent of heavy-atom derivatization. In most cases, however, the rate-limiting step is determined by the rate of chemical reaction rather than the rate of diffusion. Mass-spectrometry measurements indicated that covalent heavy-atom derivatizations of proteins in solution were also complete within minutes (Sun & Hammer, 2000). Under crystallization conditions, heavy-atom reaction rates vary depending on pH, buffer and the precipitants present in the soaking solutions. Using lysozyme and FcγRIII crystals as test cases for several heavy-atom compounds, we have demonstrated that quick soaks yield equally good or better derivatives than prolonged soaks. In these cases, the standard 1–2 d

Table 1
Data collection and phasing statistics.

Statistics for the outermost resolution shell are listed in parentheses. The outer shells of the native, quick- and long-soak TBRII data are 1.05–1.09, 1.34–1.39 and 1.18–1.22 Å, respectively. The outer shells of the native, the edge and remote wavelengths of the quick-soak NKG2D–ULBP3 data are 2.6–2.7, 3.1–3.2 and 2.9–3.0 Å, respectively.

	TBRII			NKG2D–ULBP3		
	Native	Quick soak	Long soak	Native	Quick soak	
Data collection						
Heavy atom		$HgCl_2$	$HgCl_2$		K_2PtCl_4	
Concentration		Saturated	Saturated		10 m*M*	
Soaking time		10 min	12 h		10 min	
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P4_32_12$	$P4_32_12$	
λ† (Å)	1.00920	1.00764 (p)	1.00764 (p)	0.9715	1.07212 (e)	1.0537 (r)
Unit-cell parameters (Å)						
a	35.5	35.0	34.4	62.0	62.5	62.4
b	40.6	40.5	40.4	62.0	62.5	62.4
c	76.1	76.0	76.3	237.3	238.5	238.2
Resolution (Å)	1.05	1.3	1.2	2.6	3.2	3.0
Completeness (%)	87.9	87.6	82.2	98.9	97.6	97.8
I/σ	56.5 (2.8)	27.7 (2.7)	23.0 (1.9)	15.6 (3.4)	16.2 (4.0)	21.0 (3.6)
R_{sym}	0.035 (0.49)	0.029 (0.24)	0.030 (0.29)	0.089 (0.43)	0.09 (0.49)	0.066 (0.51)
R_{iso}		0.23	0.37		0.325	0.34
Heavy-atom peak height‡						
Site 1 (σ)		26.2	17.6		12.3	15.9
Site 2 (σ)		20.6	15.8		8.5	11.0
Site 3 (σ)					7.9	10.6
Phasing method		Anomalous only			MAD	
R_{cullis}		0.45	0.48		0.71	0.75
Phasing power		3.4	3.1		1.3	1.45
FOM		0.45	0.39		0.41	

† For comparison between the quick soak and long soak, only data corresponding to the Hg absorption peak wavelength (p) are listed in the table for TBRII. Similar results are obtained for the edge and remote wavelengths. In the case of NKG2D–ULBP3, a two-wavelength MAD data was collected corresponding to the Pt L_{III} edge (e) and remote (r) wavelengths. ‡ The heavy-atom sites are shown as the peak heights (in standard deviations) from the difference Fourier maps phased with the refined TBRII or NKG2D–ULBP3 structure, respectively.

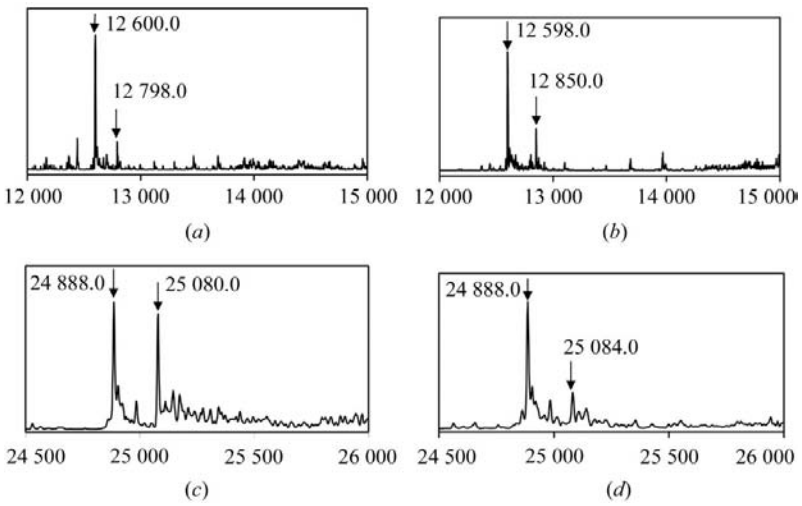

Figure 1
Mass-spectrometry measurements of the heavy-metal derivatives of TBRII and NKG2D–ULBP3. (*a*) and (*b*) show the spectra of $HgCl_2$- and TMLA-derivatized TBRII, respectively. The peaks labeled with masses of 12 598 and 12 600 Da correspond to the native TBRII, which is predicted to have a mass of 12 599 Da. The calculated masses of Hg^{2+} and $(CH_3)_3Pb^+$ are 200 and 252 Da, respectively. All observed derivative adducts have masses consistent with the predicted masses within 2 mass units. (*c*) and (*d*) are ULBP3 derivatized with K_2PtCl_4 and $KAuCl_4$, respectively. The mass of the 24 888 Da peak corresponds to that of native ULBP3 with a bound glutathione. The masses of Pt and Au are 195 and 197 Da, respectively.

soaks did not result in better heavy-atom occupancies compared with the 10 min soaks. The phasing statistics from the quick soaks were in fact better than the longer soaks, a result likely to be attributable to less non-isomorphism in the quick-soaked crystals. Here, we have applied the quick-soak method to obtain heavy-atom derivatives for *de novo* phasing of two new protein structures, the extracellular domain of a type II human transforming growth factor β (TGFβ) receptor (TBRII) and the natural killer cell receptor NKG2D in complex with its ligand ULBP3. Combined with mass-spectrometry-assisted heavy-atom screening, quick soaking has the potential to transform the conventional lengthy heavy-atom derivative-screening process into a real-time 'on-the-fly' derivatization process.

2. Materials and methods

The structure solutions of TBRII and the NKG2D–ULBP3 complex are described elsewhere (Boesen *et al.*, 2002; Radaev *et al.*, 2001). Heavy-atom derivatives were screened using electrospray ionization mass spectrometry (ESI-MS) as described previously (Sun & Hammer, 2000). Briefly, 1 µl of protein at 10 mg ml^{-1} concentration was mixed with 1 µl of 10 m*M* heavy-atom solution prepared either in water or in crystallization buffer and incubated for 30 min before mass-spectrometry analysis. Heavy-atom soaking solutions were prepared by dissolving the desired amount of heavy-atom compounds into crystallization well solutions. Preliminary heavy-atom derivative diffraction data were collected in-house using an R-AXIS IV image-plate system. The final diffraction data were collected at the X9B beamline of the NSLS (National Synchrotron Light Source) facility at Brookhaven National Laboratory. All data were processed using *HKL*2000 (Otwinowski & Minor, 1997). The heavy-atom binding sites were located using *SHELX* and *SOLVE* (Sheldrick, 1990; Terwilliger & Berendzen, 1999) and phases were determined using *CNS* version 1.0 (Brünger *et al.*, 1998).

3. Results and discussion

3.1. Phasing of the TBRII structure

TBRII is a type I transmembrane receptor ubiquitously expressed in all cells. The 14 kDa extracellular domain of TBRII binds to transforming growth factor β (TGFβ) and initiates cellular responses leading to growth arrest. The soluble portion of TBRII has been expressed and crystallized (Boesen *et al.*, 2000). Several heavy-atom compounds, including $HgCl_2$, K_2PtCl_4, $KAu(CN)_2$, $KAuCl_4$, ethylmercuric thiosalicylate (EMTS), trimethyllead acetate (TMLA) and lead (II) acetate, were screened for potential derivatizations using electrospray ionization mass spectrometry (ESI–MS). Of these, $HgCl_2$ and $PbAc_2$ generated covalent adducts with TBRII (Fig. 1). Crystals of TBRII were derivatized by soaking in well solutions containing saturated $HgCl_2$ ($\sim$10 mM) for 10 min. A three-wavelength MAD data set was collected at the Hg L_{III} absorption edge.

Both the native and $HgCl_2$-derivative crystals diffracted to high resolution (1.1 and 1.3 Å, respectively; Table 1). An anomalous difference Patterson map calculated using the quick-soak data clearly revealed Hg binding (Fig. 2). Two major Hg-binding sites were identified using *SHELX* and *SOLVE* and refined to yield a phasing power of 3.4 (Fig. 3*a*). A section of the MAD-phased best Fourier map is shown in Fig. 4(*a*).

For comparison, MAD data were also collected from a crystal derivatized for 12 h using a heavy-atom soaking solution identical to that of the quick soak (Table 1). Overall, the phasing statistics are similar between the quick and 12 h soaks, illustrating the effectiveness of the quick soak in phasing. Moreover, the quick soak yielded a derivative with slightly higher Hg occupancies and better values for the Cullis R factor, phasing power and the figure of merit than the 12 h soak. Interestingly, the calculated isomorphous R factor (R_{iso}) of the quick-soak derivative (0.23) is lower than that of the longer soak (0.37), indicating an increased crystal non-isomorphism as the result of prolonged soaking. This is also reflected in a 1.1 Å change in the unit-cell parameter a from the 12 h soak compared with a 0.5 Å change in a from the 10 min soak.

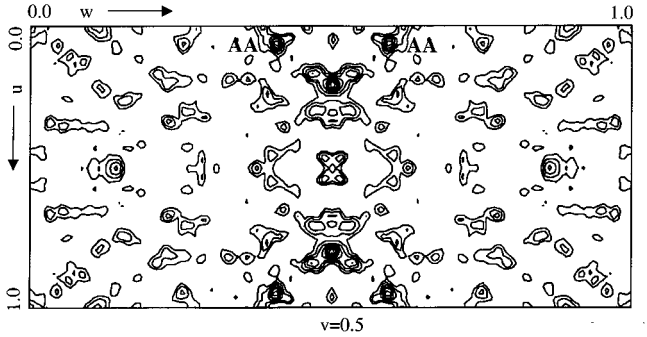

Figure 2
Anomalous difference Patterson map from the TBRII crystal derivatized by a quick soak in saturated $HgCl_2$ contoured at 1σ. The v = 0.5 Harker section is shown, with the Harker peaks generated from the first Hg-binding site labeled AA.

Since phases derived from isomorphous replacement ($F_{PH} - F_P$) terms are affected by crystal non-isomorphism between a derivative and its native, they are often not coherent with phases derived from anomalous and multi-wavelength components. Attempts to combine these phases often yield worse electron-density maps than MAD phasing alone without the contribution from isomorphous replacement phases. To test whether a quick soak minimizes crystal non-isomorphism compared with a longer soak, electron-density maps were calculated using SIRAS (single isomorphous replacement and anomalous scattering) phases derived from both the 10 min and the 12 h soaked derivatives of TBRII. When compared with Fourier maps calculated from the MAD phases, a better SIRAS map was obtained from the phases of the shorter soak, whereas a worse SIRAS map was obtained from the phases of longer soak (Fig. 4). This difference in SIRAS map quality between the soaks demonstrates the advantage of a quick soak in reducing the effect of crystal non-isomorphism.

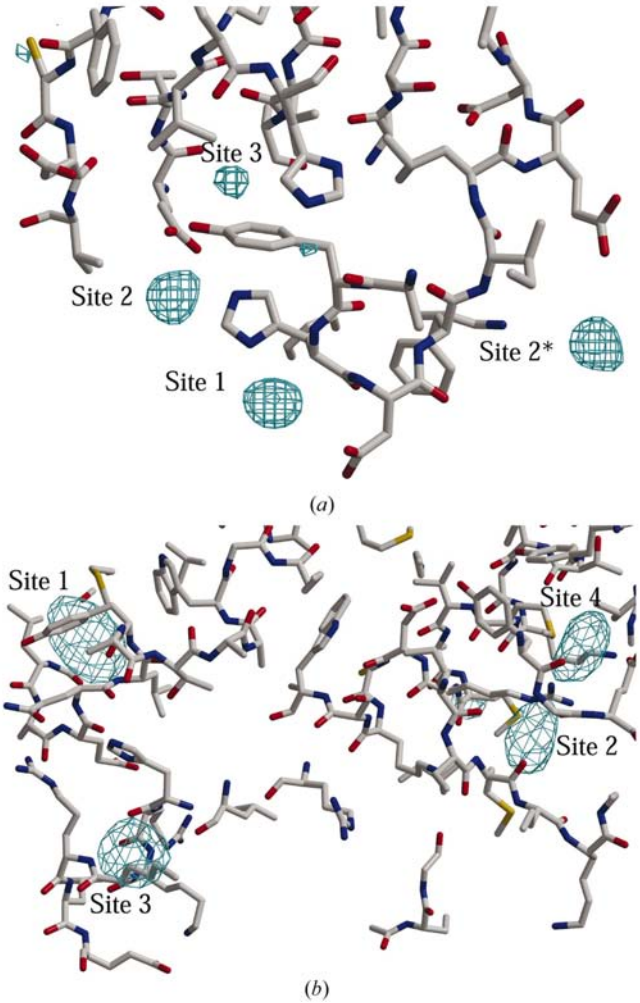

Figure 3
Difference Fourier maps showing (*a*) the three Hg-binding sites on TBRII and (*b*) the four Pt-binding sites on ULBP3. The maps are contoured at 5σ with 1σ increments. * indicates symmetry-related heavy-atom sites.

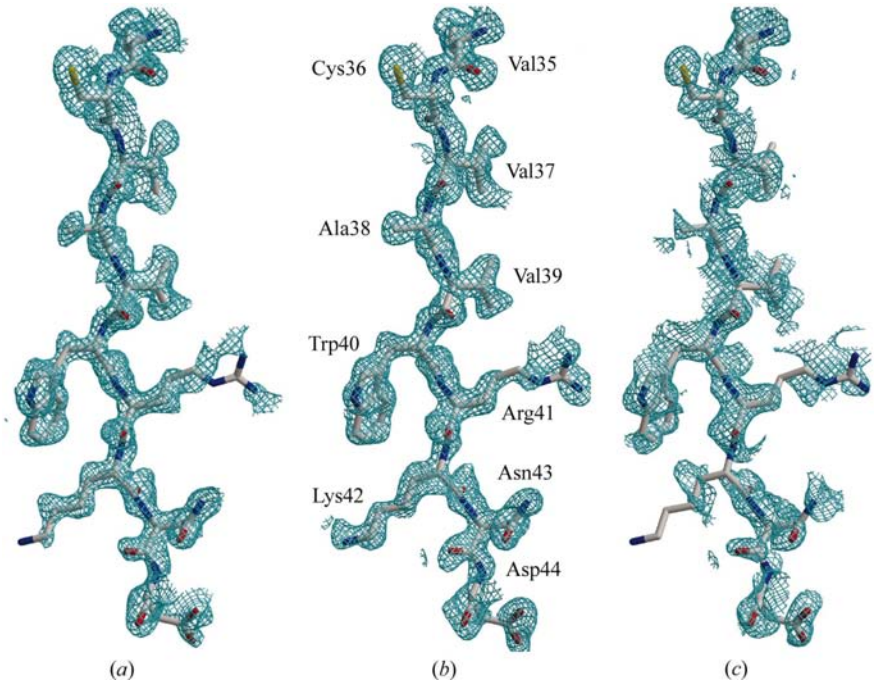

Figure 4
Experimental electron-density maps of TBRII phased with $HgCl_2$ dervivatives. (*a*) A region of the MAD phased electron-density map contoured at 1σ with the corresponding refined model. (*b*) SIRAS map produced by a 10 min quick soak; (*c*) SIRAS map resulting from the long 12 h soak.

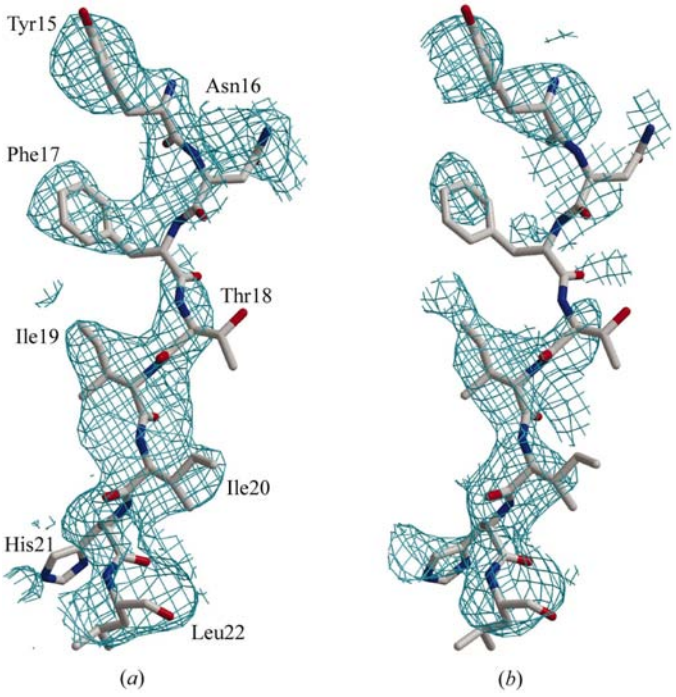

Figure 5
Experimental electron-density maps of the NKG2D–ULPB3 complex phased with a quick-soaked K_2PtCl_4 derivative. (*a*) Electron-density map generated from MAD and SIR combined phases contoured at 1σ displaying a β-strand from ULBP3. (*b*) Electron-density map produced from MAD phases alone showing the same region as (*a*).

The solution of the TBRII structure provides an example of *de novo* phasing using a heavy-atom derivative generated from a 10 min quick soak. The quality of this quick-soak $HgCl_2$ derivative is evident from its Cullis *R* factor and phasing power as well as the experimental electron-density map (Table 1, Fig. 4).

3.2. Phasing of the NKG2D–ULBP3 crystal

NKG2D is a 14 kDa C-type lectin-like receptor expressed on the surface of natural killer cells and certain T cells. ULBP3 is a 24 kDa class I major histocompatibility (MHC) antigen-like molecule and a ligand of NKG2D. The binding of ULBP3 activates natural killer cells and triggers NKG2D-mediated lysis of ULBP3 bearing target cells (Cosman *et al.*, 2001). The complex between the dimer of NKG2D and ULBP3 has been crystallized in the tetragonal space group $P4_32_12$ and diffracts to 2.6 Å resolution (Radaev *et al.*, 2001). The receptor half of the complex was solved by molecular replacement using the structure of a homologous murine NKG2D receptor. The ligand ULBP3, however, was not visible in the NKG2D phased electron-density map. To provide phases for the ligand part of the structure, several heavy-atom compounds, including K_2PtCl_4, $KAuBr_4$, $KAuCl_4$, $HgCl_2$, $HgBr_2$, TMLA, $SmCl_3$ and mercuric fluorescein, were screened using ESI–MS for potential derivatives of ULBP3. Among them, K_2PtCl_4, $KAuBr_4$ and $KAuCl_4$ showed heavy-atom adducts in mass-spectrometry analysis (Fig. 1). All attempts to soak NKG2D–ULBP3 crystals for 24 h in solutions containing 1 m*M* of the designated heavy-atom compound, however, resulted in partial lattice disorder and loss of diffraction beyond 6 Å. In contrast, a quick soak of the crystals in 10 m*M* K_2PtCl_4 for 10 min resulted in no visual deterioration of the diffraction. A two-wavelength MAD data set was collected at the Pt L_{III} absorption edge and a remote wavelength (Table 1).

A total of four Pt-binding sites were found in the asymmetric unit of the K_2PtCl_4 derivative using a difference Fourier map phased using the NKG2D half of the complex (Fig. 3*b*). Refinement of the Pt sites resulted in a phasing power of 1.3 for the Pt derivative and an overall figure-of-merit of 0.41 for all phased reflections. To examine the possibility of combining the phases derived from the isomorphous term with those derived from the anomalous scattering, best Fourier density maps were calculated using either MAD alone or SIR–MAD combined phases. The combination of SIR and MAD phases resulted in a better electron-density map than that calculated from the MAD phases alone (Fig. 5).

It is worth emphasizing that only the quick soak resulted in a usable heavy-atom derivative for the phasing of the NKG2D–ULBP3 structure and that all longer soakings resulted in overall crystal lattice disorder. Unlike the TBRII crystals, which diffracted to 1.1 Å resolution, the crystals of NKG2D–ULBP3 diffracted to a moderate 2.6 Å resolution. This suggests that brief soaks may be advantageous over conventional longer soaks, especially for low-resolution diffracting crystals that have weaker crystal lattice packing and are more subject to disruptions by heavy-atom soaks.

The ability to obtain useful derivatives for both TBRII and NKG2D–ULBP3 crystals demonstrates the validity of using the quick-soak method to derivatize protein crystals. In both cases, prolonged soaks are not necessary to achieve the desired heavy-atom occupancies. In practice, the success of this method depends on the reaction rate of heavy-atom compounds being such that faster reacting heavy atoms will be more suitable for quick soaks. This rate, which varies depending on conditions such as buffer, pH and other chemical groups present in crystallization solutions, can be readily accessed in solution by mass spectrometry (Sun *et al.*, 2002). In doing so, one maximizes heavy-atom occupancies for the benefit of phasing, while minimizing the soaking time in order to preserve the lattice order. A limitation to mass-spectrometry-assisted heavy-atom screening is its ability to detect electrostatically bound heavy-atom adducts, although bound ions have been observed as protein adducts in mass spectrometry. As an alternative to mass-spectroscopy optimization, one can start with a 2 h soaking at near-saturation heavy-atom compound concentration and adjust the soaking time and heavy-atom concentration by diffraction screening.

In some cases, particularly in lower resolution crystals, prolonged soaking leads to profound lattice disorder, as seen in overnight-soaked NKG2D–ULBP3 crystals. The quick-soak method is therefore the only viable method for generating potential derivatives of NKG2D–ULBP3 crystals.

3.3. Potential application to structural genomics with 'on-the-fly' heavy-atom phasing

In the era of structural genomics, high-throughput structure determination demands more efficient methods to generate heavy-atom phasing than the conventional slow soaking method. Currently, the selenomethionine and halide-soaking methods are the preferred derivatization methods for applications in structural genomics. While both methods avoid the lengthy heavy-atom derivative screening process, each has its own limitations. The selenomethionine method is primarily applied to bacterially expressed recombinant proteins, although limited success has been achieved in yeast or baculovirus expression systems. The halide-soaking method has been successfully applied to some small proteins with molecular weights less than 20 kDa and to crystals that diffract to high resolution. The quick-soak method shortens heavy-atom derivatizations from the conventional time frame of days to minutes. When combined with mass-spectrometry heavy-atom screening, it offers a real-time on-the-fly method of obtaining heavy-atom derivatives for phasing.

In summary, we demonstrate that good *de novo* protein phases can be obtained from derivatives prepared by the quick-soak method. For real-time heavy-atom derivatizations, we recommend preliminary screening of potential derivatives using solution reactions with mass spectrometry under crystallization buffer conditions. In principle, one can also optimize heavy-atom concentration and soaking time using mass spectrometry. Once potential derivative compounds are identified by mass spectrometry, the quick-soak method should then be applied to derivatize the crystals. We recommend heavy-atom concentrations between 10 and 50 m*M* or at saturation for less soluble compounds with soaking times between 10 min and 2 h.

We thank C. Hammer for providing the mass-spectrometry measurements, and C. Boesen and J. Johnson for assistance in synchrotron data collection. This work is supported by the funding for NIAID intramural research.

References

Boesen, C. C., Motyka, S. A., Patamawenu, A. & Sun, P. D. (2000). *Protein Expr. Purif.* **20**, 98–104.
Boesen, C. C., Radaev, S., Motyka, S. A., Patamawenu, A. & Sun, P. D. (2002). In the press.
Brünger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, N., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Cosman, D., Mullberg, J., Sutherland, C. L., Chin, W., Armitage, R., Fanslow, W., Kubin, M. & Chalupny, N. J. (2001). *Immunity*, **14**, 123–133.
Nagem, R. A., Dauter, Z. & Polikarpov, I. (2001). *Acta Cryst.* D**57**, 996–1002.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Radaev, S., Rostro, B., Brooks, A. G., Colonna, M. & Sun, P. D. (2001). *Immunity*, **15**, 1039–1049.
Sheldrick, G. (1990). *Acta Cryst.* A**46**, 467–473.
Sun, P. D. & Hammer, C. H. (2000). *Acta Cryst.* D**56**, 161–168.
Sun, P. D., Radaev, S. & Kattah, M. (2002) *Acta Cryst.* D**58**, 1092–1098.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.