Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Protein crystallization by rational mutagenesis of surface residues: Lys to Ala mutations promote crystallization of RhoGDI

Kenton L. Longenecker, Sarah M. Garrard, Peter J. Sheffield and Zygmunt S. Derewenda*

Department of Molecular Physiology and Biological Physics, University of Virginia, PO Box 800736, Charlottesville, VA 22908-0736, USA

Correspondence e-mail: zsd4n@virginia.edu

Received 27 November 2000
Accepted 14 February 2001

PDB References: Δ23 K135,138,141A mutant, 1fst; Δ66 K135,138,141A;L196F mutant, 1fso; Δ66 K113A mutant, 1ft0; Δ66 K141A mutant, 1ft3.

Crystallization is a unique process that occurs at the expense of entropy, including the conformational entropy of surface residues, which become ordered in crystal lattices during formation of crystal contacts. It could therefore be argued that epitopes free of amino acids with high conformational entropy are more thermodynamically favorable for crystal formation. For a protein recalcitrant to crystallization, mutation of such surface amino acids to residues with no conformational entropy might lead to enhancement of crystallization. This paper reports the results of experiments with an important cytosolic regulator of GTPases, human RhoGDI, in which lysine residues were systematically mutated to alanines. Single and multiple mutations were introduced into two different variants of RhoGDI, NΔ23 and NΔ66, in which the first 23 and 66 residues, respectively, were removed by recombinant methods. In total, 13 single and multiple mutants were prepared and assessed for crystallization and all were shown to crystallize using the Hampton Research Crystal Screens I and II, in contrast to wild-type NΔ23 and NΔ66 RhoGDI which did not crystallize. Four crystal structures were solved (the triple mutants NΔ23:K135,138,141A and NΔ66:K135,138,141A, and two single mutants NΔ66:K113A and NΔ66:K141A) and in three cases the crystal contacts of the new lattices were found precisely at the sites of mutations. These results support the notion that it is, in principle, possible to rationally design mutations which systematically enhance proteins' ability to crystallize.

1. Introduction

Established protein crystallization methods rely on systematic screening for suitable conditions, including various precipitants, pH values, additives, temperature *etc*. In many cases, thousands of conditions are tested during screening; new crystallization robots are expected to set up over 100 000 samples daily (Stevens, 2000). However, the effectiveness of crystallization experiments is not directly proportional to the number of conditions tested and current attempts at global enhancement of structure-determination methodology (structural genomics) urgently require improvements in crystallization effectiveness.

It could be argued that the protein itself, rather than the precipitating agent, may be considered to be the most important variable in the screening process. Some proteins (*e.g.* lysozyme, insulin, hemoglobin) exhibit surface characteristics which lead to rapid crystallization under a variety of conditions, while other proteins have never been crystallized. Clearly, those proteins have surface properties that are not

favorable for the formation of crystal contacts. The possibility of using the protein as one of the variables in the process of screening for crystals has been suggested in the past. As early as 1972, it was proposed that if crystals of a target protein cannot be grown, homologous proteins from other species should be considered (Campbell *et al.*, 1972). In this way, these authors crystallized nearly every enzyme in the glycolytic pathway, albeit from different organisms. Although this approach is still often used, there is no way to predict if a protein from one species is more likely to crystallize than that from another. Another approach, which has so far enjoyed less popularity, is to mutate surface residues in search of a variant more susceptible to crystallization. One of the first examples of such protein engineering to promote crystallization was the work of Lawson and colleagues, who mutated human ferritin to generate crystal contacts analogous to those in the rat isoform (Lawson *et al.*, 1991). This was followed in 1992 by a seminal study which showed that even single-site amino-acid substitutions in thymidylate synthase substantially affected the crystallization of the protein (McElroy *et al.*, 1992). Since then, several examples of successful crystallization of mutants in lieu of wild-type proteins have been reported. During the studies of the HIV type 1 integrase, mutants in which hydrophobic residues were systematically replaced by lysines or alanines led to the identification of a single-site mutant, F185K, which gave crystals used in the successful structure analysis (Dyda *et al.*, 1994; Jenkins *et al.*, 1995). It has been reported that different mutants of T4 lysozyme can be crystallized in as many as 25 different forms (Zhang *et al.*, 1995). The structural characterization of the *obese* protein leptin E100 (Zhang *et al.*, 1997) was possible because the W100→E mutation allowed the preparation of good crystals. Histidine ammonia lyase was successfully crystallized by exchange of a surface cysteine residue to an alanine (Schwede *et al.*, 1999). Very recently, D'Arcy and colleagues at Hoffman–La Roche used mutagenesis of surface residues in the DNA gyrase B subunit to study the effect of surface substitutions on the ability of the protein to crystallize and obtained new crystal forms (D'Arcy *et al.*, 1999). In many cases, serendipity was a factor: the first crystals of the chaperone GroEL were obtained after errors were introduced inadvertently at the PCR stage (Horwich, 2000).

In spite of these results, the concept of crystal engineering has suffered from the lack of a rational approach: mutations were clearly recognized as a means to change proteins' solubility, but no rational protocol was ever designed to crystallize a new protein by surface mutagenesis. We hypothesized that mutagenesis of surface residues such as Lys and Glu to Ala or other smaller amino acids might systematically improve protein crystallization by creating epitopes more favorable for the formation of crystal contacts with negligible loss of conformational entropy. In this paper, we describe our initial testing of this approach on the human guanine nucleotide dissociation inhibitor RhoGDI. This physiologically important regulator of Rho-family GTPases has been previously shown to be recalcitrant to crystallization. RhoGDI is a single-chain protein made up of 204 residues, with a functional moiety restricted to the fragment 26–204 (Platko *et al.*, 1995). In the absence of its target, Rho GTPase, residues 23–68 remain disordered (Gosser *et al.*, 1997; Keep *et al.*, 1997). Full-length protein and N-terminally truncated (Δ23 and Δ66) variants are all easily expressed in *Escherichia coli* using various fusion constructs (Sheffield *et al.*, 1999). A tryptic fragment (59–204) has been crystallized and its structure has been reported (Keep *et al.*, 1997) revealing an immunoglobulin-like fold; no successful crystallization of either the Δ23 or Δ66 variants has ever been described. The protein contains 14 lysines (10%) and 12 glutamates (8%) in the globular domain (67–210), making it a good candidate to test the Lys→Ala (K2A) concept (Fig. 1). We show that K2A mutations notably improve the success rate in crystallization of RhoGDI using commercially available Crystal Screen kits from Hampton Research. We also show that the integrity of crystal lattices in the

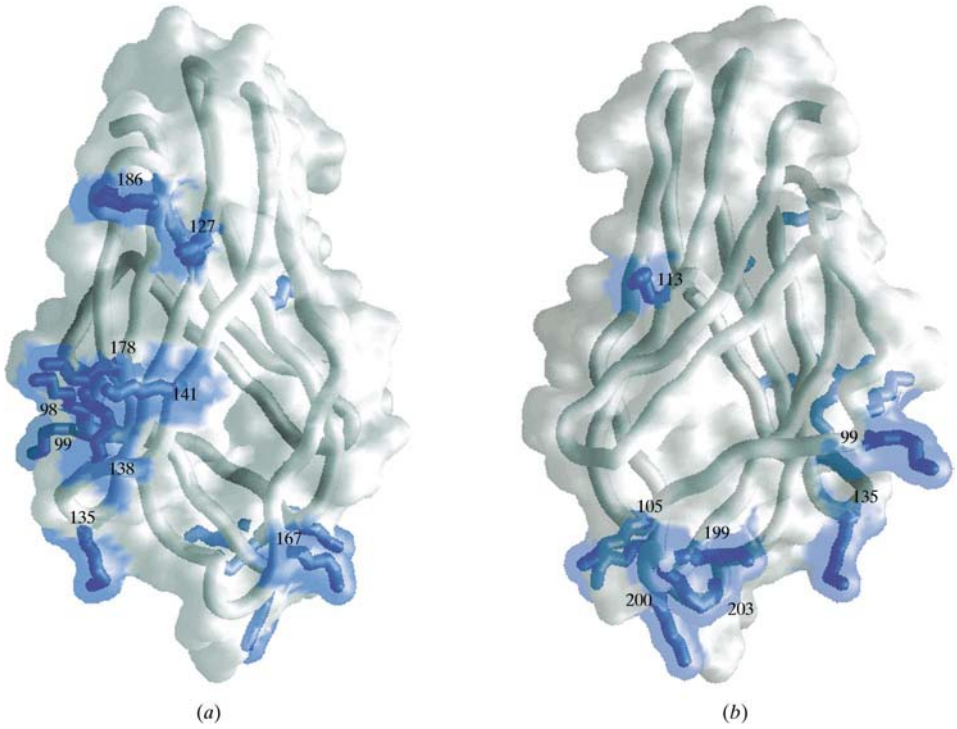

Figure 1
Overall structure of the ordered C-terminal domain of RhoGDI. Panels (*a*) and (*b*) show front and back perspectives of the β-sandwich topology of the immunoglobulin-like fold as viewed through a transparent rendering of the molecular surface. The 14 lysine residues in this domain are highlighted in blue.

Table 1
Primers used in the design of mutants.

Mutation	Background	Primer sequence
K105A	Δ23N	GCAGTCGTTTGTcCTGgcGGAGGGTGTGGAG-TACCGG
K113A	Δ23N and Δ66N	GGAGTACCGGATAgcgATaTCTTTCCGGG
K127A	Δ23N	CCGAGAGATAGTGagCGGCATGgcGTACATCCAGC
K141A	Δ23N and Δ66N	GGCGTCAAaATcGAtgcGACTGACTACATGGTAGGC
K98,99,105A	Δ23N and Δ66N	CGGGCGACCTGGAGAGCTTCgcGgcG-CAaagcTTTGTGCTGgcGGAGGGT
K135,138,141A	Δ23N and Δ66N	CCAGCATACGTACAGGgcAGGCGTCgcGATTGACgc-GACTGACTACATGG
K98,99,105,113A	Δ23N	Two-step mutagenesis using the primer for 113A and the template of 98,99,105
K98,99,105,141A	Δ23N	Two-step mutagenesis using the primer for 141A and the template for 98,11,105

mutants is maintained predominantly by intermolecular contacts involving sites of mutations.

2. Materials and methods

2.1. RhoGDI plasmids and mutagenesis

Using pGEX2T1::RhoGDI wild-type plasmid as template DNA, new ATG start codons and *Nco*I sites were introduced into RhoGDI at positions 23 and 66 by the polymerase chain reaction. Two mutagenic primers (Δ23N, 5′-GACCCCAC-CATGGCCAACGTCGTGGTGACTGG; Δ66N, 5′-GCAGA CCCCATGGTCCCCAACGTCG-3′) and a pGEX-specific primer (5′-CACCGTCATCACCGAAACGCGCGAGGC-3′) were used to amplify the truncated RhoGDI genes. The subsequent products were digested with *Nco*I and *Eco*RI and subcloned into pGST-Parallel1 (Sheffield *et al.*, 1999) similarly digested. Positive clones, pGST::RhoGDIΔ23N and pGST::RhoGDIΔ66N, were verified by DNA sequencing. All mutations were introduced into genes coding for Δ66 and Δ23 RhoGDI in fusion with GST, with an rTEV proteinase cleavage site engineered for effective removal of the GST tag, as described elsewhere (Sheffield *et al.*, 1999). The QuikChange protocol has been used throughout, modified in our laboratory to reduce the cost. Primers were purchased from Life Technologies and used after dissolving in distilled water. Single and triple mutants described in our study were prepared using single mutagenic primers because of the proximity of the mutated sites; the triple mutants were prepared by mutagenesis of the plasmids carrying the triple mutation in a two-stage reaction. A complete list of primers and mutations is shown in Table 1.

2.2. Protein expression and purification

Expression of fusion proteins was carried out in 1–4 l of XL1 Blue cultures in LB medium incubated at 310 K until an OD_{600} of 0.4–0.6. Expression was then induced by the addition of IPTG (final concentration 500 μ*M*). Cultures were grown for an additional 4 h at 310 K. Cells were harvested by centrifugation, resuspended in buffer *A* (50 m*M* Tris–HCl pH 8.5, 300 m*M* NaCl; ∼5 ml per gram of cell pellet) and lysed using a French press. The soluble protein fraction was recovered by centrifugation. GST fusion proteins were isolated by binding to glutathione Sepharose columns (Pharmacia) pre-equilibrated with 100 ml of buffer *A*. The soluble protein fraction was passed through the column twice at 277 K and the columns were washed with 2 l of buffer *A* at 277 K. GST fusion protein was eluted in 20 ml of buffer containing 10 m*M* reduced glutathione pH 8.5. The GST tag was removed by treatment with 100 units of rTEV proteinase (Life Technologies), typically for 48 h at 283 K during dialysis against 2 l of buffer *A*, to remove the glutathione before being reapplied to a glutathione Sepharose column to remove uncut protein as well as the contaminating GST tag. This procedure removed over 90% of the GST protein; the residual GST was removed by size-exclusion chromatography (Superdex 75). Residual rTEV was also removed by passing the purified protein through a mini-spin Ni^{2+} column (Qiagen). Proteins were analyzed by SDS–PAGE on 12% polyacrylamide gels in a Laemmli buffer system. Protein concentrations were measured using the Bradford assay (BioRad).

2.3. Crystallization

Proteins were crystallized at 293 K by vapor diffusion in VDX plates using Crystal Screens I and II (Hampton Research). Typically, 2 μl protein solution was mixed with 2 μl reservoir and suspended on a cover slip over a reservoir containing 0.5 ml precipitant. The concentration of the protein samples varied between 7 and 15 mg ml^{-1}. The Δ23 K135,138,141A mutant crystallized in 12–14% PEG 3400, 5% 2-propanol and 100 m*M* HEPES pH 7.0. The Δ66 triple mutant crystallized in 44–48% saturated ammonium sulfate and 100 m*M* MES pH 6.0. The Δ66 single mutants crystallized similarly in 50% saturated ammonium sulfate, 200 m*M* sodium/potassium tartrate and 100 m*M* sodium citrate pH 5.6.

2.4. X-ray data collection

Single crystals of the mutants were transferred to a stabilizing solution containing the reservoir components and a cryoprotectant and were then flash-frozen by plunging into liquid nitrogen. The cryosolution for the Δ23 triple mutant contained no 2-propanol but contained 30% PEG 400, whereas the cryoprotectant for the other crystals was 25–30% glycerol. X-ray data were collected on frozen crystals at beamline X11 at the EMBL Outstation Hamburg facility on a single-chip MAR Research CCD area detector and were processed with the *DENZO*/*SCALEPACK* set of programs (Otwinowski & Minor, 1997). Data for the Δ66 triple mutant were similarly collected from a single crystal at beamline X9B of the NSLS (Brookhaven National Laboratory) and were processed as above. Intensities were converted to structure-

Table 2
X-ray data statistics.

Values in parentheses are for the highest resolution shells, including data between d_{min} and 2.78, 2.07, 2.66 or 2.85 Å for each mutant, respectively.

	GDI-Δ23 K135,138,141A	GDI-Δ66 K135,138, 141A;L196F	GDI-Δ66 K113A	GDI-Δ66 K141A
Space group	*R*3	$P3_221$	*R*32	*R*32
Unit-cell parameters (Å)				
a, *b*	125.2	76.3	128.5	129.6
c	82.7	88.6	164.3	164.8
d_{min} (Å)	2.7	2.0	2.6	2.8
No. of observations	49713	83504	73487	57299
No. unique	13241	20565	15110	13249
Completeness (%)	99.6	99.7	92.8	99.7
$I/\sigma(I)$	17.9 (3.7)	13.7 (4.8)	16.8 (2.7)	11.0 (1.9)
R_{sym}† (%)	4.8 (26.7)	2.9 (30)	5.8 (33)	9.4 (68)
Beamline	EMBL X11	NSLS X9B	EMBL X11	EMBL X11

† $R_{sym} = \sum |I - \langle I \rangle| / \sum I$, where I is the integrated intensity for a particular reflection.

Table 3
Structure refinement statistics.

	GDI-Δ23 K135,138,141A	GDI-Δ66 K135,138, 141A;L196F	GDI-Δ66 K113A	GDI-Δ66 K141A
Resolution range (Å)	20–2.7	20–2.0	20–2.6	20–2.8
R factor (%)	21.3	21.6	22.6	21.6
R_{free}† (%)	26.1	23.2	26.0	26.0
Reflections for R_{free}	1310	1050	1001	869
Protein atoms	2274	1106	2150	2155
Solvent molecules	40	199	56	36
Mean *B* factor (Å^2)	47.2	41.6	52.8	52.5
R.m.s.d. bonds (Å)	0.007	0.006	0.007	0.007
Rm.s.d. angles (°)	1.37	1.39	1.48	1.44

† R_{free} is calculated as the conventional crystallographic *R* factor, but using a randomly selected subset of reflections not incorporated into the refinement procedures.

factor magnitudes using the program *TRUNCATE* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). Crystallographic details for all four sets are shown in Table 2.

2.5. Structure solution and refinement

Structures were solved by molecular replacement using the 2.5 Å resolution model of RhoGDI (chain *A* of PDB code 1rho) with the *AMoRe* program (Navaza, 1994). Molecular-replacement calculations on data from the Δ23 triple mutant yielded two independent solutions better than any others, which upon proper placement relative to one another yielded a correlation coefficient of 62 and an *R* factor of 37% for data in the resolution range 12–4.0 Å. Similar calculations on the Δ66 triple-mutant data (12–3.5 Å) yielded a single solution with a correlation coefficient of 49.8 and an *R* factor of 43%, whereas the values for the next best wrong solution were 21 and 53%, respectively. Molecular replacement on the Δ66 single-mutant data sets yielded only one clear individual solution, but after fixing this initial solution the second solution became apparent upon subsequent searches. The combined solutions had a correlation coefficient of 49.7 and an *R* factor of 41.3% for data in the resolution range 10–4.0 Å.

The atomic models were refined against the structure-factor magnitudes using maximum-likelihood calculations implemented in *CNS* (Brunger *et al.*, 1998); visual inspection of electron-density maps and rebuilding were performed using *O* (Jones *et al.*, 1991). Solvent was identified automatically in difference maps and added manually for the lower resolution structures. The solvent structure for the Δ66 triple mutant was built using *ARP* in conjunction with *REFMAC*, both from the *CCP*4 suite (Murshudov *et al.*, 1997; Perrakis *et al.*, 1999). The stereochemistry of the models was evaluated with *PROCHECK* (Laskowski *et al.*, 1993) and each was of acceptable quality. The results of the refinement are tabulated in Table 3. Figures were produced using *GRASP* (Nicholls *et al.*, 1991), *BOBSCRIPT* (Esnouf, 1997) and *MOLSCRIPT* (Kraulis, 1991).

Table 4
Hits obtained in crystallization screens.

CS1, Crystal Screen I; CS2, Crystal Screen II; the numbers correspond to the solution number in the Hampton Research Screens (components noted below)†. The bold numbers illustrate conditions which yielded crystals used in X-ray diffraction studies.

pGST-GDI Δ23N	Crystals	Diffraction	CS1	CS2
K105A	Yes	—	4, 39 and 32	25, 32 and 14
K113A	Yes	—	32	15
K127A	Yes	—	16 and 17	23 and 25
K141A	Yes	—	4 and 32	14, 15 and 25
K98,99,105A	Yes	Yes	15	—
K135,138,141A	Yes	Yes	**41**	—
pGST-GDI Δ66N				
K113A	Yes	Yes	32	**14**
K141A	Yes	—	—	**14**
K98,99,105A	Yes	Yes	4, 6, 9, 15, 16, 17 and 33	20, 23, 25, 26, 32 and 42
K135,138,141A	Yes	Yes	—	**32**
pGST-GDI Δ23N 9899105				
K113A	Yes	—	39	25, 32
K141A	Yes	—	6, 15, 17	26

† In CS1: 4, 2.0 *M* ammonium sulfate (AS), 0.1 *M* Tris–HCl pH 8.5; 6, 30%(*w*/*v*) PEG 4000, 0.1 *M* Tris–HCl pH 8.5, 0.2 *M* $MgCl_2$ hexahydrate; 9, 30% PEG 4000, 0.1 *M* sodium citrate pH 5.6, 0.2 *M* ammonium acetate; 15, 30% PEG 8000, 0.1 *M* sodium cacodylate pH 6.5, 0.2 *M* AS; 16, 1.5 *M* lithium sulfate monohydrate, 0.1 *M* Na HEPES pH 7.5; 17, 30% PEG 4000, 0.1 *M* Tris–HCl pH 8.5, 0.2 *M* lithium sulfate monohydrate; 32, 2.0 *M* AS; 33, 4.0 *M* sodium formate; 39, 2%(*v*/*v*) PEG 400, 2.0 *M* AS, 0.1 *M* Na HEPES pH 7.5; 41, 10%(*v*/*v*) propanol, 20%(*w*/*v*) PEG 4000, 0.1 *M* Na HEPES pH 7.5. In CS2: 14, 2.0 *M* AS, 0.1 *M* trisodium citrate dihydrate pH 5.6, 0.2 *M* potassium/sodium tartrate tetrahydrate; 15, lithium sulfate monohydrate, buffer as above, 0.5 *M* AS; 20, 1.6 *M* magnesium sulfate heptahydrate, 0.1 *M* MES pH 6.5; 23, 10%(*v*/*v*) dioxane, buffer as above, 1.6 *M* AS; 25, 1.8 *M* AS, 0.1 M MES pH 6.5; 26, 30% PEG MME 5000, buffer as above, 0.2 *M* AS; 32, 1.6 *M* AS, 0.1 *M* HEPES pH 7.5, 0.1 *M* NaCl; 42, 12% glycerol, 0.1 *M* Tris pH 8.5, 1.5 *M* AS.

3. Results and discussion

3.1. K2A mutations enhance crystallization of RhoGDI

Neither of the two N-terminally truncated variants of human RhoGDI (*i.e.* NΔ23 and NΔ66) crystallized from any of the Crystal Screen I and II solutions. Using the Δ23 construct of RhoGDI, we prepared four single mutants, two

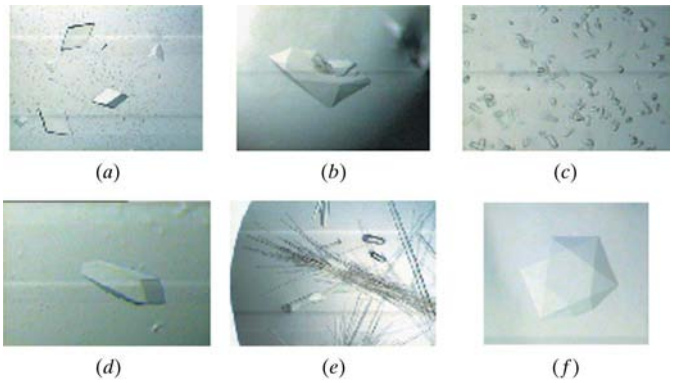

Figure 2
Microphotographs of selected crystals optimized from initial crystal hits. (*a*) Δ23 K135,138,141A triple mutant, CS1, 41. (*b*) Δ66 K113A single mutant, CS2, 14. (*c*) Δ66 K98,99,105A triple mutant, CS1, 33. (*d*) Δ23 K98,99,105A triple mutant, CS1, 15. (*e*) Δ66 K135,138,141A triple mutant, CS2, 32. (*f*) Δ66 K141A single mutant, CS2, 14.

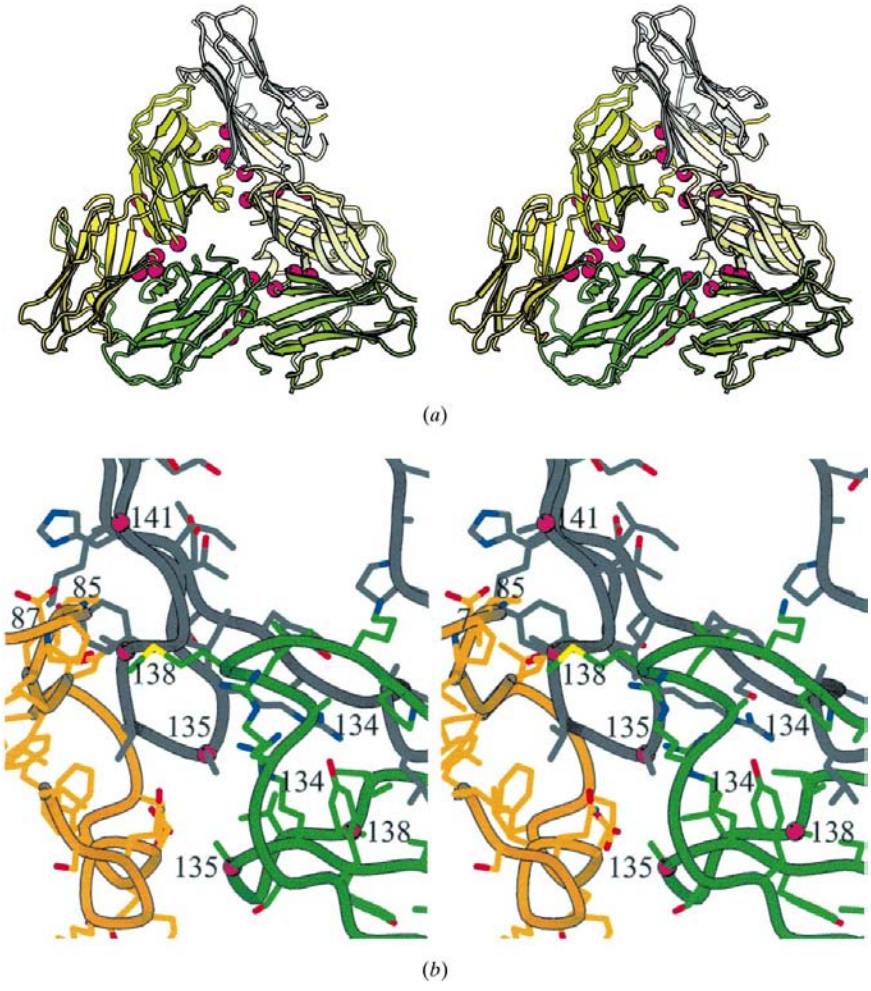

Figure 3
Stereoview of the molecular packing in crystals of the Δ23 triple mutant. (*a*) Several molecules are shown from a view looking down the crystallographic threefold axis. The location of K2A point mutations are highlighted with magenta balls. (*b*) A close-up view reveals molecular interactions between molecule *A* (grey) and molecule *B* (green) of the asymmetric unit as well as another molecule (orange) related by crystallographic symmetry.

triple mutants and two quadruple K2A mutants (Table 1). No special rationale was used in the selection of sites; in particular, triple mutants were designed based on the proximity of selected sites to allow mutagenesis using a single set of primers. All mutants gave 'hits' in either Crystal Screen I or II; five preparations (K105A, K127A, K141A and both quadruple mutants) gave three to five hits (Fig. 2; Table 4). It must be remembered that the Δ23 construct contains a disordered N-terminal fragment over 40 residues long. Thus, at least in the case of RhoGDI, surface engineering appears to outweigh the negative impact of partial disorder. The Δ66 construct, containing only the immunoglobulin-like domain, was used to generate two single mutants (K113A, K141A) and two triple mutants, both analogous to those tested in the Δ23 construct. Again, all samples gave hits in the Hampton screens, with the K98,99,105A triple mutant giving 13 hits in total (Table 4). In practically all cases crystallization occurred quickly, within hours or at most a few days of setting up the screens. Crystals were obtained from a variety of solutions, but most hits were found in ammonium sulfate, with or without additives and with a range of buffers. Crystals also grew from PEG of higher molecular weights, lithium sulfate and magnesium sulfate. In one case, crystallization was observed in 4.0 *M* sodium formate solution.

Crystals of six of the mutants were assessed by X-ray diffraction using in-house facilities. Two, Δ23:K98,99,105A and Δ66:K98,99,105A, showed large unit cells which were found to be impractical for fast investigation by molecular-replacement procedures. The remaining four exhibited acceptable unit cells and were pursued further (Tables 2 and 3). The choice was somewhat arbitrary, since probably any of the initial 'hits' could have been optimized to yield large single crystals for X-ray diffraction. Two crystals selected for structure determination contain the triple mutation K135,138,141A, albeit in two different constructs with N-termini of various lengths. All four crystals were used to collect data to ∼2.5 Å using synchrotron radiation. It was subsequently found that the Δ66:K135,138,141A mutant crystals diffracted significantly better when an additional mutation L196F, originally designed to probe the functionality of the hydrophobic pocket of RhoGDI in an unrelated study, was introduced. This did not change the crystallization habit (*i.e.* the crystals are isomorphous with those of Δ66: K135,138,141A), but increased the resolution to 2.0 Å. Thus, the discussion of crystal contacts in this form applies directly to the triple mutant.

3.2. The crystal structure of the Δ23 K135,138,141A mutant

The crystal structure of the Δ23:K135,138,141A triple mutant was solved by molecular replacement and refined at 2.7 Å resolution to a conventional R factor of 21.3% ($R_{\rm free}$ of 26.1%). This crystal form (space group $R3$) contains two protein molecules (denoted A and B) related by non-crystallographic twofold symmetry. The solvent content is fairly high, reflected in a Matthews coefficient of 3.1 Å^3 Da^{-1}. The non-crystallographic dimer is formed by a tightly packed symmetric interface that is centered upon Arg134 and has numerous van der Waals interactions, including the Ala135 introduced in place of a lysine (Fig. 3*a*). Adjacent to this interface, we find another close contact between the A molecule and another B molecule, mediated by the epitope containing mutated sites 138 and 141. The nature of this interface is decidedly hydrophobic, with Pro85 of molecule B lodged between the C$^{\beta}$ methyl groups of Ala138 and Ala141. The contact is further stabilized by the antiparallel association of two short stretches of main chain, so that two strong main-chain–main-chain hydrogen bonds are made: B87 CO to NH of A138 (2.97 Å) and A136 CO to NH of B89 (2.96 Å). It is quite clear that this close interface, involving primarily main-chain atoms, could not form with the wild-type lysines in positions 135, 138 and 141 (Fig. 3*b*).

In addition to the contacts involving the mutated epitope, molecule B makes an additional extensive contact with an adjacent pair of symmetry-related A molecules mediated by the N-terminal fragment encompassing residues 56–65. This fragment, not present in the Δ66 construct, wedges between the two A molecules in such a way that it forms an additional strand along one of the β-sheets in RhoGDI. This interesting and extensive contact is obviously insufficient to generate crystals of the wild-type Δ23 RhoGDI; it appears that the above-described contacts involving the mutated 135, 138 and 141 sites play a critical role.

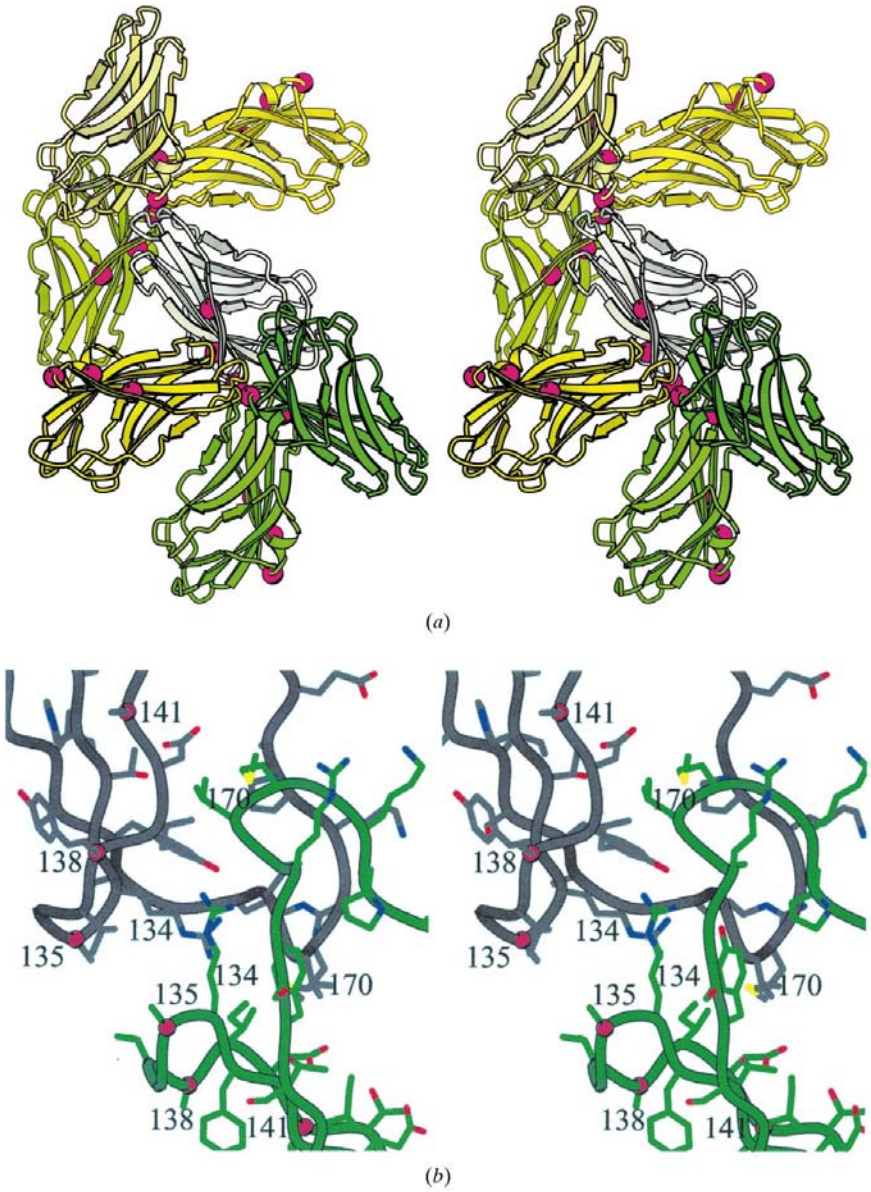

Figure 4
Stereoview of the molecular packing in crystals of the Δ66 triple mutant. (*a*) Several molecules are shown from a view looking down the crystallographic threefold screw axis. The location of K2A point mutations are highlighted with magenta balls. (*b*) A close-up view reveals molecular interactions between two molecules related by crystallographic symmetry. This view is reminiscent of the packing of non-crystallographic dimers in the other crystal forms. For clarity, this view does not show the packing of two other crystallographically related molecules, which are also in close proximity and are apparent in Fig. 4(*a*).

3.3. The structure of the Δ66 K135,138,141A mutant

In spite of the same set of surface mutations as the above sample, the shorter Δ66 variant forms crystals exhibiting different symmetry, *i.e.* $P3_221$. Although considerations based on the Matthews coefficient ($V_{\rm M}$ of 4.6 Å^3 Da^{-1}, assuming one molecule) initially indicated that two molecules may occupy the asymmetric unit in these crystals, subsequent molecular-replacement calculations persistently gave only one solution. This solution was subjected to refinement using data to 2.0 Å resolution and the process converged well with a conventional R factor of 21.6% ($R_{\rm free}$ = 23.2%). The packing diagram (Fig. 4*a*) shows how an elongated molecule of RhoGDI forms a stable crystal lattice mediated through crystal contacts involving residues at two opposite ends.

The loop responsible for much of the crystal-forming interface is again precisely

the loop containing the triple mutation K135,138,141A (Fig. 4*b*). This interface brings into proximity two symmetry-related pairs of molecules, so that two molecules contribute the 135,138,141 loop while the other two are involved *via* the topologically opposite end. The resulting extensive interface is populated with small mostly hydrophobic residues with low conformational entropy, *e.g.* Val137, Val67, Ala153, Pro151, Ile122 and notably Ala135 and Ala138. There are few direct interactions and only three hydrogen bonds, all involving main-chain atoms: 139 CO to 67 NH (3.2 Å), 120 CO to 174 NH (2.8 Å) and 167 CO to 167 NH (3.2 Å) from another molecule. Lack of a larger number of direct bonds confers a character of non-specificity on this interface. The void spaces in this interface are filled with ordered water molecules forming local networks of hydrogen bonds.

3.4. The structures of the Δ66 K113A and K141A single-site mutants

Both single mutants K113A and K141A crystallize in the same crystal form (space group *R*32) and like the previous crystals forms show a relatively high solvent content, with a Matthews coefficient of 4.1 Å^3 Da^{-1}. The asymmetric unit contains a dimer which is conspicuously similar to the dimer found in the Δ23 K135,138,141A mutant in that the two molecules insert the hydrophobic residues of the 169–170 loop into the hydrophobic cavity of the adjacent molecule (Fig. 5*a*). The rest of the dimer is largely free of crystal contacts, with the exception of the surface around Ala113, the site of mutation in one of the mutants. The limited contact in this area involves several residues of mixed character, involving Ala113, Val78 and a few neighboring residues from the *A* molecule packed against Pro85, Glu87 and Val78 in the symmetry-related *B* molecule (Fig. 5*b*). There are no immediately obvious direct hydrogen bonds at this interface, but water does fill the void spaces. In the K141A mutant, the contact retains its integrity, even though Lys113 now fills some space at the interface. Interestingly, the mutation site is distant from any crystal contacts and it is difficult to rationalize why the K141A mutant yielded this crystal form whereas other preparations of Δ66 truncation mutants did not.

3.5. Overview of the structural data

The choice of Lys→Ala (K2A) mutations in this pilot study was dictated by many factors. Lysines are predominantly located on the surface, with 68% exposed, 26% partly exposed and only 6% buried (Baud & Karlin, 1999). In the case of RhoGDI, all of the lysines are either completely or partially exposed on the surface. On average, Lys is the most solvent-exposed residue in proteins. Statistically, these amino acids constitute 5.8% of the total amino-acid content in proteins, ranging from zero in rare and small acidic proteins to over 10% in some microtubule-associated proteins. It is important in this context to realise that even in a protein with average lysine content (5.8%) these amino acids constitute a large fraction of the surface area because nearly all are solvent accessible. Thus, even in a protein with an unknown fold, lysines provide markers for identifying a solvent-accessible site with a better than 90% success rate. Intuitively, we assume that the high conformational entropy of solvent-exposed lysines, ∼8 kJ mol^{-1} (Avbelj & Fele, 1998), is likely to impair the formation of protein–protein contacts. This is supported by a recent study which shows that lysines constitute only 5.4% of interface surface in oligomeric proteins, compared with 14.9% of the total solvent-accessible surface (Conte *et al.*, 1999).

We hypothesized that mutating Lys to Ala may be effective in enhancing crystallizability, as alanine is evenly distributed

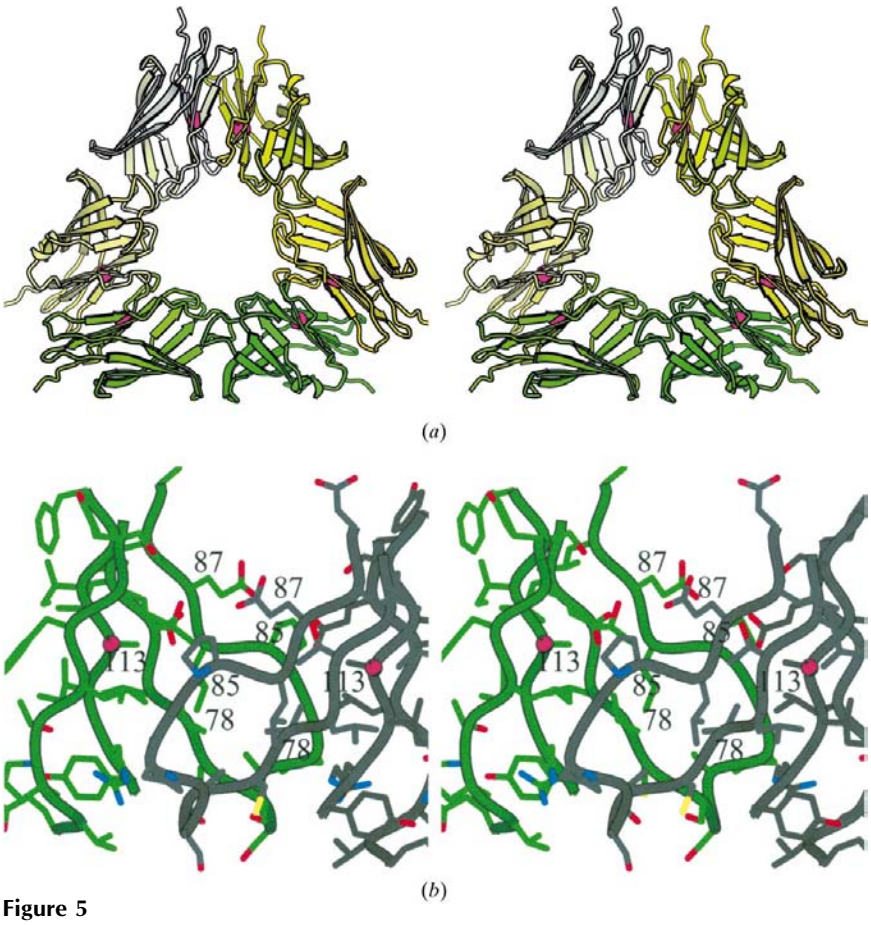

Figure 5
Stereoview of the molecular packing in crystals of the Δ66 K113A single mutant. (*a*) An overall view is shown looking down the crystallographic threefold axis. The location of residue 113 is highlighted with magenta balls. (*b*) A close-up view reveals molecular interactions between molecule *A* (grey) and molecule *B* (green) of the asymmetric unit.

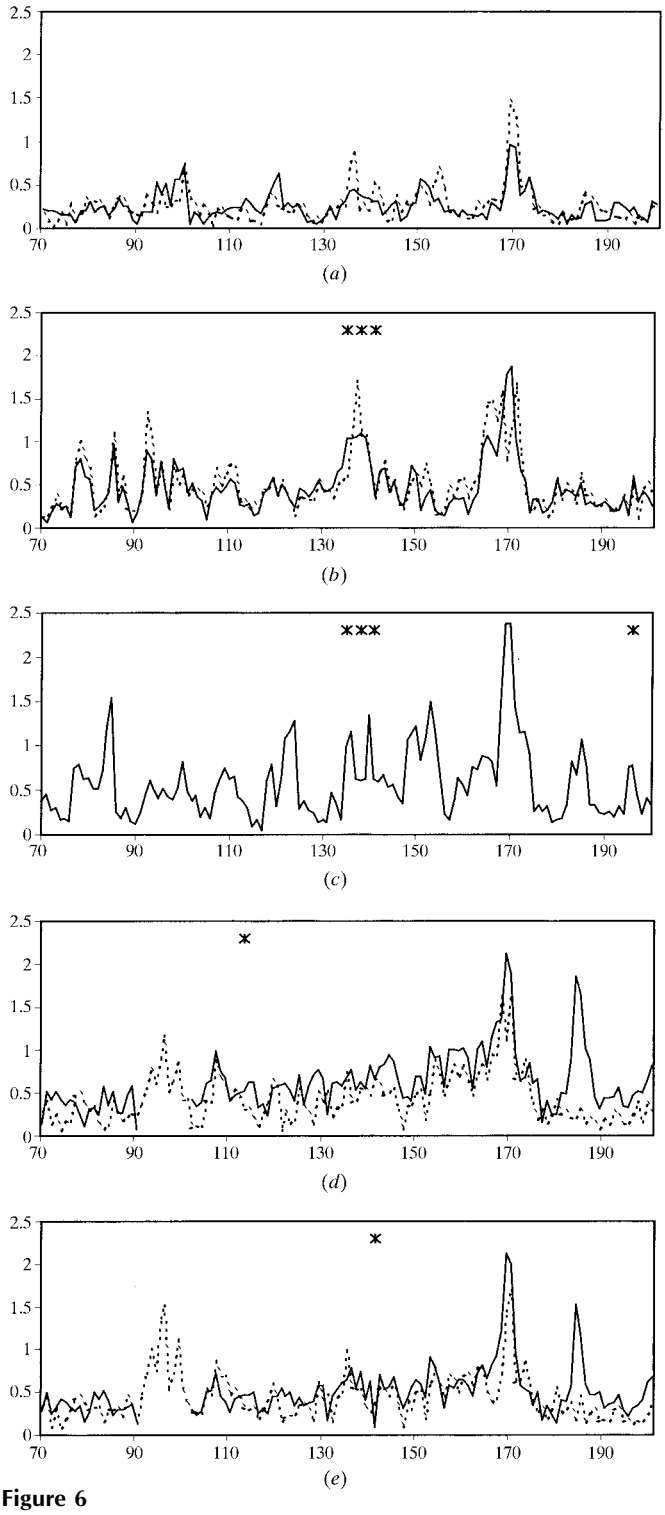

Figure 6
Comparison of the differences between structures of RhoGDI. Plots are shown for the distances (Å) between the C^{α} positions of aligned RhoGDI structures as a function of residue position. Positions of point mutations are highlighted with an asterisk. (*a*) Difference plots are shown for the three molecules in the asymmetric unit of the native structure (PDB code 1rho): molecule *A* compared with molecule *B* (solid) or molecule C (dashed) of the asymmetric unit. (*b*)–(*e*) Difference plots are shown for comparisons between molecule *A* of 1rho and the mutant structures reported in this manuscript, using solid lines for molecule *A* and dashed lines for molecule *B* where appropriate. The mutants are as follows: (*b*) Δ23 triple mutant, (*c*) Δ66 triple mutant, (*d*) Δ66 K113A single mutant and (*e*) Δ66 K141A single mutant.

in the interior, on the surface and at the interface in oligomeric proteins (Conte *et al.*, 1999) and has no conformational entropy. With the exception of the K141A mutant, the structural data strongly supports our hypothesis that reduction of surface conformational entropy by mutagenesis can lead to the creation of epitopes thermodynamically suitable for the formation of crystal lattices. In the two triple-mutant structures, the sites of mutation are clearly involved in the crystal contacts and seem to be directly responsible for the crystallization habit of the protein. The crystal contacts involve numerous interactions between main-chain atoms, including hydrogen bonds between main-chain carbonyls and amides. It is not clear to what extent the K113A mutation was helpful in promoting the *R*32 crystal form. Although residue 113 is located at an intermolecular crystal contact, the remotely positoned K141A mutation leads to a similar packing arrangement with minimal changes to accommodate the lysine at position 113. While all K2A mutants proved more susceptible to crystallization efforts, we infer from the current data that multiple mutations are perhaps more likely to generate useful contiguous surface epitopes allowing the preparation of crystals in a variety of lattice arrangements.

The structures solved in our study indicate that K2A mutations did not cause any significant structural perturbations in RhoGDI. We have analyzed the structures of all mutants by overlapping them using a least-squares algorithm and calculating plots of C^{α} differences compared with the original search model (Fig. 6). In general terms, we note no significant conformational changes arising from mutations, even in the mutants containing triple K2A substitutions located in close proximity along the polypeptide chain. Some deviations are observed in the C^{α} plots, but they are smaller than the differences observed for the 170 loop which is also involved in crystal contacts. There are only minor conformational changes in the side chains surrounding the mutation sites and they appear to be caused by crystal packing interactions rather than the mutations themselves.

A comparison of the C^{α} coordinates in the two molecules of the triple Δ23 mutants with the search model from the PDB shows that the mean r.m.s. for the C^{α} positions is about 0.7 Å, consistent with the limited resolution of both X-ray studies. The largest deviation of nearly 2 Å is found in the flexible loop that contains residue 170, a pattern that we also observe in other structures. The fragment containing the mutated sites shows slightly higher deviation than average and the two molecules differ somewhat in the extent of changes. Since this site mediates crystal contacts, it is difficult to judge if the observed differences are a consequence of mutagenesis or of the crystal contact. A comparison of C^{α} coordinates of the 2.0 Å resolution structure of the triple Δ66 mutants with the search model reveals an overall r.m.s. of approximately 0.75 Å, with few C^{α} atoms differing by more than 1.0 Å. Many of the differences are likely to stem from the relative inaccuracy of the 2.5 Å resolution search model. The C^{α} atoms of the mutated sites are within this range, while the largest discrepancy is again observed for the loop that contains residue 170. The plots of C^{α} differences for the two single-site mutants

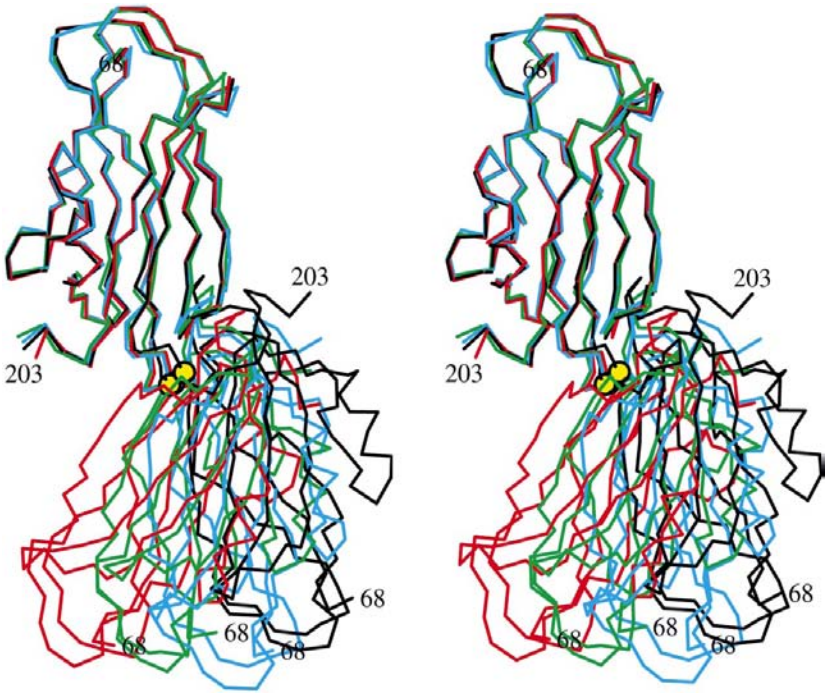

Figure 7
Similarity of the molecular interactions between various crystal forms of RhoGDI. A stereoview highlights the propensity of RhoGDI to form dimers, an arrangement that is evident in all crystal forms. Overlaying the structures (C traces) of one of the molecules from the dimer shows the relative placement of the other molecule. The crystals represented are of the wild-type (black), the Δ23 triple mutant (red), the Δ66 quadruple mutant (cyan) and the Δ66 K113A single mutant (green). The location of residue 170 is highlighted in the aligned molecules with yellow spheres.

reveal no major changes when compared with the search model, again with the notable exception of the 170 loop, which deviates by over 2 Å.

Parenthetically, we note that the RhoGDI molecules have a propensity to form non-crystallographic dimers in which the hydrophobic entrances to the central cavity face each other and in all crystal lattices a contact of this nature is evident (Fig. 7). The 170 loop, whose apex sequence is GMLA, is invariably in close contact with the residues lining the entrance in the adjacent molecule and its conformation is altered as the mutual disposition of the two RhoGDI molecules varies from dimer to dimer. This explains the observed conformational flexibility of the 170 loop seen in the different crystal lattices.

4. Conclusions

Our results strongly suggest that, at least in the model system of RhoGDI, K2A mutagenesis led to improved crystallization results. Out of the 13 various mutants, all gave hits and several gave diffraction-grade crystals in the first screens. Furthermore, the structures of four mutants reveal that crystal contacts are in fact mediated by epitopes involving the mutated sites. These results strongly suggest a causal relationship between the Lys→Ala mutations and enhanced crystallization of the RhoGDI mutants compared with the wild type.

To date, site-directed mutagenesis has not been used widely to enhance crystallization, probably owing to the lack of established protocols that would provide guidance to the type of mutation that is required. The approach we propose is based on a hypothesis that residues with high conformational entropy may impede crystallization. Other factors may clearly contribute to the enhanced crystallizability of K2A mutants (*e.g.* removal of potential steric hindrance, alteration of electrostatic charge, the hydration shell *etc.*), but whatever the rationale, the important outcome of our study is the demonstration of a causal relationship between K2A mutations and crystallization. To establish whether that is indeed the case, we obviously need to test other types of mutations (involving, for example, Glu and Gln) and other protein systems. This work is under way.

This work was funded by NIH grants HL48807 and NS36267. We would like to thank Drs Zbigniew Dauter, Jia Li and Yancho Devedjiev for help in collecting X-ray diffraction data at the synchrotron beamlines. We also thank Ms Ruby Hsiu for assistance in protein sample preparation.

References

Avbelj, F. & Fele, L. (1998). *J. Mol. Biol.* **279**, 665–684.
Baud, F. & Karlin, S. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 12494–12499.
Brunger, 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.
Campbell, J. W., Duee, E., Hodgson, G., Mercer, W. D., Stammers, D. K., Wendell, P. L., Muirhead, H. & Watson, H. C. (1972). *Cold Spring Harbor Symp. Quant. Biol.* **36**, 165–170.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Conte, L. L., Chothia, C. & Janin, J. (1999). *J. Mol. Biol.* **285**, 2177–2198.
D'Arcy, A., Stihle, M., Kostrewa, D. & Dale, G. (1999). *Acta Cryst.* D**55**, 1623–1625.
Dyda, F., Hickman, A. B., Jenkins, T. M., Engelman, A., Craigie, R. & Davies, D. R. (1994). *Science*, **266,** 1981–1986.
Esnouf, R. M. (1997). *J. Mol. Graph.* **15**, 132–143.
Gosser, Y. Q., Nomanbhoy, T. K., Aghazadeh, B., Manor, D., Combs, C., Cerione, R. A. & Rosen, M. K. (1997). *Nature (London)*, **387**, 814–819.
Horwich, A. (2000). *Nature Struct. Biol.* **7**, 269–270.
Jenkins, T. M., Hickman, A. B., Dyda, F., Ghirlando, R., Davies, D. R. & Craigie, R. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 6057–6061.
Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst* A**47**, 110–119.

Keep, N. H., Barnes, M., Barsukov, I., Badii, R., Lian, L. Y., Segal, A. W., Moody, P. C. E. & Roberts, G. C. K. (1997). *Structure*, **5**, 623–633.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 282–291.
Lawson, D. M., Artymiuk, P. J., Yewdall, S. J., Smith, J. M., Livingstone, J. C., Treffry, A., Luzzago, A., Levi, S., Arosio, P., Cesareni, G., Thomas, C. D., Shaw, W. V. & Harrison, P. M. (1991). *Nature (London)*, **349**, 541–544.
McElroy, H. H., Sisson, G. W., Schottlin, W. E., Aust, R. M. & Villafranca, J. E. (1992). *J. Cryst. Growth*, **122**, 265–272.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Nicholls, A., Sharp, K. & Honig, B. (1991). *Proteins Struct. Funct. Genet.* **11**, 281–296.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.
Platko, J. V., Leonard, D. A., Adra, C. N., Shaw, R. J., Cerione, R. A. & Lim, B. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 2974–2978.
Schwede, T. F., Badeker, M., Langer, M., Retey, J. & Schulz, G. E. (1999). *Protein Eng.* **12**, 151–153.
Sheffield, P., Garrard, S. & Derewenda, Z. (1999). *Protein Expr. Purif.* **15**, 34–39.
Stevens, R. C. (2000). *Curr. Opin. Struct. Biol.* **10**, 558–563.
Zhang, F., Basinski, M. B., Beals, J. M., Briggs, S. L., Churgay, L. M., Clawson, D. K., DiMarchi, R. D., Furman, T. C., Hale, J. E., Hsiung, H. M., Schoner, B. E., Smith, D. P., Zhang, X. Y., Wery, J. P. & Schevitz, R. W. (1997). *Nature (London)*, **387**, 206–209.
Zhang, X. J., Wozniak, J. A. & Matthews, B. W. (1995). *J. Mol. Biol.* **250**, 527–552.