Published in final edited form as:

Cryst Growth Des. 2009 August 5; 9(8): 3570–3574. doi:10.1021/cg9003142.

Protein Crystal Engineering of YpAC-IV using the Strategy of Excess Charge Reduction

D. Travis Gallagher,

Biochemistry Division, National Institute of Standards and Technology, Gaithersburg, MD, 20899-8312

N. Natasha Smith,

Biochemistry Division, National Institute of Standards and Technology, Gaithersburg, MD, 20899-8312

Sook-Kyung Kim,

^aCenter for Bioanalysis, Korea Research Institute for Standards and Science, Yuseong-Gu, Republic of Korea

Howard Robinson, and

^bBiology Department, Brookhaven National Laboratory, Upton, New York, 11973

Prasad T. Reddy

Biochemistry Division, National Institute of Standards and Technology, Gaithersburg, MD, 20899-8312

Abstract

The class IV adenylyl cyclase from *Yersinia pestis* has been engineered by site-specific mutagenesis to facilitate crystallization at neutral pH. The wild-type enzyme crystallized only below pH 5, consistent with the observation of a carboxyl-carboxylate H bond in a crystal contact in the refined structure 2FJT. Based on that unliganded structure at 1.9 Å resolution, two different approaches were tested with the goal of producing a higher-pH crystal needed for inhibitor complexation and mechanistic studies. In one approach, Asp 19, which forms the growth-limiting dicarboxyl contact in wild-type triclinic crystals, was modified to Ala and Asn in hopes of relieving the acid-dependence of that crystal form. In the other approach, wild-type residues Met 18, Glu 25, and Asp 55 were (individually) changed to lysine to reduce the protein's excess negative charge in hopes of enabling growth of new, higher-pH forms. These 3 sites were selected based on their high solvent exposure and lack of intraprotein interactions. The D19A and D19N mutants had reduced solubility and did not crystallize. The other 3 mutants all crystallized, producing several new forms at neutral pH. One of these forms, with the D55K mutant, enabled a product complex at 1.6 Å resolution, structure 3GHX. This structure shows why the new crystal form required the mutation in order to grow at neutral pH. This approach could be useful in other cases where excess negative charge inhibits the crystallization of low-pI proteins.

Keywords

crystal growth; crystal contact; crystal engineering; isoelectric point; mutation; structure; x-ray diffraction

*Corresponding author: 9600 Gudelsky Drive, Rockville, MD, 20850, Tel. 240-314-6204; Fax 301-975-5449; E-mail gallaghe@umbi.umd.edu.

1. Introduction

When a wild-type protein fails to produce adequate crystals despite extensive trials, mutation of its surface residues is often a useful approach¹. When the structure is unknown, residues predicted to lie on the surface may be mutated according to any of several strategies that generally seek, for high-solubility proteins, to reduce surface entropy², and for low-solubility cases, to improve solubility^{3,4}. When the structure is known, mutations can be precisely targeted to crystal contacts in order to improve the diffraction or pH profile of an existing crystal form⁵. Such an approach was used to change a low-pH requiring, platelike crystal form of subtilisin into one with more isometric morphology and a broader pH profile⁶. Alternatively, mutations can be directed to disrupt an existing contact, or simply to alter the known surface (perhaps by entropy reduction or solubility-enhancement), in hopes of producing a more suitable new crystal form.

In general, the structures and properties of new crystal contacts and forms that result from these mutations are not amenable to prediction, even when the protein structure is known. However, some statistical data are accumulating regarding the usefulness of various mutational substitutions, largely based on engineered mutations that promoted crystallization in prior studies. The entropy-reducing mutations Lys → Ala and Glu → Ala have been found effective in many cases⁷. Reduction of the protein's conformational entropy is also the rationale of strategies that identify and remove disordered termini and mobile loops⁸. The replacement of Lys by Arg has been evaluated as a strategy⁹. Most of these approaches apply in cases where protein solubility is good, as they tend to reduce solubility. When proteins are marginally soluble, fusions with highly soluble 'carrier proteins' have been described, and some mutational strategies to improve solubility and crystallizability have been found effective^{10, 11}. Replacement of DTNB-accessible cysteines by serines was found to increase solubility and crystallizability dramatically in chorismate lyase¹².

In this paper we begin with the triclinic crystal form of the wild-type *Yersinia pestis* adenylate cyclase IV (YpAC-IV, cyaB, predicted pI~5.0). Figure 1 shows triclinic YpAC-IV crystals, which grow as intricate clusters of thin plates, with individual crystallites rarely over 10 μm thick¹³. This crystal form grows only in the pH range 4.5 (below which, precipitation) to 5.0 (above which, no crystals). These limitations prevented using these crystals to solve the structure or for substrate-analog complexation (for which low pH is unfavorable). Nevertheless this form produced one 25 μm thick crystal that furnished the data for structure 2FJT¹⁴. The difficulties with this form were circumvented by altering the expression construct from His-tagged-thrombin-removed, to native, thus removing a 3-residue artifact from the N-terminus. This native construct produced crystals in a new orthorhombic form that diffracted only to 3 Å but enabled structure determination. Because both crystal forms of wild-type YpAC-IV grew only below pH 5, active-site ligand complexation was effectively prevented, and we sought an approach to producing crystals at higher pH, either a new form, or through a modification of the high-diffracting triclinic form to enable it to exist at higher pH.

In order to produce an active-site complex crystal structure for mechanistic study, we engineered the YpAC-IV protein surface at several sites. Examination of the wild-type structure 2FJT revealed a crystal contact apposing two Asp 19 residues as the likely cause for the thin morphology and pH-limited growth of the triclinic morphodrom, as previously shown in a subtilisin variant⁶, suggesting mutations at this site. Another mutant, D55K, crystallized in a new form at pH 6.5 that was suitable for soaking and complexation with the product cAMP. The resulting complex yielded 1.6 Å diffraction data, providing a detailed view of product binding in the active site. In the resulting structure, a crystal contact featuring the mutant D55K residue includes two carboxylate groups and thus appears to explain why wild-type crystals do not grow in this form.

2. Materials and Methods

2.1 Mutagenesis and protein preparation

Mutants of *Yersinia pestis* AC-IV were prepared using the QuikChange II XL Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA). Mutagenic oligonucleotides were obtained from Qiagen (Valencia, CA) and were purified on an HPLC column and phosphorylated with polynucleotide kinase. The mutagenesis efficiency was $\approx 90\%$. The template plasmid DNA for mutagenesis was the pET11a recombinant. The YpAC2 coding region was sequenced to confirm the mutations. For purification of the protein, culture was scaled up to 1 L of cells. Cells (1.5 g wet weight) were suspended in 40 mL of 50 mM Tris-HCl, pH: 7.5, 1 mM DTT, 1 mM EDTA, and 1 mM PMSF (lysis buffer). Cell free extract was prepared by passing the cell suspension through a French press twice and centrifugation at $100\,000 \times g$ for 1 h. The supernatant was loaded on a 40 mL DE52 column, and eluted with a linear NaCl gradient generated from 200 mL of the lysis buffer and 200 mL of the lysis buffer containing 0.5 M NaCl. Fractions corresponding to the expressed protein, as judged by the Coomassie stain on SDS-PAGE, were concentrated and further purified on a 500 mL Sephadex-G75 superfine column (2.6 by 100 cm) equilibrated with the lysis buffer supplemented with 100 mM NaCl. Again, the purest protein fractions corresponding to the expressed protein were pooled and concentrated. Final step of purification was achieved by passing the concentrate from the G-75 column through 5 mL of Affigel Blue resin. YpAC2 was unbound to the resin whereas contaminating proteins were bound to the resin.

2.2 Crystallization and structure determination

AC-IV mutants were prepared for crystallization by concentration to 20 mg/mL (0.5 mM dimer) in a buffer comprising 20 mM Tris-HCl, pH 8.0, 1 mM DTT. Initial screening for crystal growth by vapor diffusion was attempted with various commercial screens from both Hampton Research and Emerald Biosystems, at both room temperature (RT, $\approx 24^\circ\text{C}$) and 4°C , using hanging drops. The D55K protein crystallized in several different morphologies with maximum dimensions less than 20 microns, and in one large platelike habit at RT from Hampton Research's crystal screen 2, reagent #26, which is 0.2 M ammonium sulfate, 30 % (w/v) PEG MME 5K, 0.1 M MES, pH 6.5. This condition was optimized to 0.1 M salt and 25 % PEG and these crystals were used for soaking with substrate analogs, resulting in several active-site complex structures at about 1.7 Å resolution.

Crystals grown at pH 6.5 were migrated to pH 7.0 (to enhance nucleotide binding and to approach the pH of optimum activity, ≈ 8.5) by incubating the coverslips (with crystal drops) over wells at the higher pH, and then soaked for 5 minutes in a modified mother liquor featuring 30 % PEG, pH 7.1, 20 mM cAMP, 15 mM MnCl_2 , all at RT. After this soak, crystals were soaked for 3 second in a cryoprotectant solution resembling the ddATP soak but containing 15 % glycerol, and then frozen by immersion in liquid nitrogen. Diffraction data to 1.6 Å resolution were collected at NSLS beamline X29A (Brookhaven National Labs, NY, USA). The structure was solved by the molecular replacement program PHASER¹⁵ using the wild-type structure 2FJT.

2.3 Refinement and structure analysis

The refinement program REFMAC¹⁶ and the molecular graphics program XFIT¹⁷ were used to adjust the model to improve its agreement with diffraction data while maintaining good geometry. Complete diffraction statistics including the overall quality indicator $R_{\text{free}} = 0.25$ are available along with the coordinates, from the Protein Data Bank under accession code 3GHX. A description of the refined cAMP complex structure along with other active site complexes, activity measurements and mechanistic analysis, is in preparation. Program

XFIT¹⁷ was also used to inspect the initial 2FJT structure in order to identify residues for mutagenesis, and to analyze crystal packing interactions. Asp 19 was observed to form a dicarboxyl contact (Figures 2,3) and was changed to Ala and Asn in hopes of mitigating the effects of that contact while preserving the crystal form. Met 18 was identified as an unusual exposed methionine, while Glu 25 and Asp 55 were identified as fully exposed negative charges with no obvious intraprotein interactions. In these latter three cases, the residues were changed to lysine to maximize solubility while reducing the protein's excess negative charge. Figure 4 shows the four mutation sites in relation to the overall dimer structure.

3. Results

Inspection of the 1.9 Å resolution wild-type structure 2FJT reveals that residue Asp 19 is involved in an unusual crystal contact, with key implications for the growth of this triclinic form. The packing is shown in Figure 2. Each Asp 19 side chain carboxylate in the crystal forms a close H-bond with another Asp 19 side chain, on a different dimer, adjacent in the (001) direction. Such a carboxylate-carboxylate H-bond requires that one of the groups is protonated, and this occurs preferentially at low pH. This contact forms a Periodic Bond Chain^{18,19} in the *c* direction, and this PBC would be expected to confer pH-sensitivity to this direction of growth. The interaction links Asp 19 of the A subunit with the same residue in the B subunit of the adjacent dimer (Figure 3). The oxygen-to-oxygen distance is 2.6 Å. This structural finding provides a molecular level explanation for the observed slow growth of the (001) face (the crystals grow as very thin plates, Figure 1), and for the crystals' inability to grow above pH 5.0¹³. It follows that mutagenesis of Asp 19 to Asn might enable this form to grow thicker, and at higher pH. A similar finding involving the protein subtilisin s88 showed that the relevant D → N mutation improved crystal morphology and removed acid-dependence⁶. We therefore made the D19N and D19A mutations in the hopes of maintaining the triclinic form, but producing crystals of thicker morphology over a wider range of pH values.

The selection of other surface residues to mutate was guided by the following considerations. We hypothesized that the tendency of the enzyme to crystallize only at low pH was related to its low isoelectric point (~5) and correspondingly high excess negative charge (this property likely also explains the protein's tendency to be precipitated by stoichiometric levels of divalent cations), and therefore sought to mutate the protein to increase its pI. This would be accomplished by changing selected Asp and Glu residues to Lys. Lys is the residue of choice in this case because of its positive charge and its tendency to promote solubility. Although the entropy change is expected to be unfavorable (especially for Asp → Lys), the more balanced surface charge is expected to compensate by enabling more crystal-contact salt links. Selection of Asp and Glu residues was also weighted toward those in relatively convex locations since they are more likely to contact other molecules. Finally, highly hydrophobic residues that are surface exposed were considered good candidates for mutation to Lys, both to increase solubility and to increase pI. Inspection of the protein surface in the 2FJT structure led to the identification of Met 18 as an unusual exposed methionine. Also the residues Glu 25 and Asp 55 were identified as highly exposed candidates in areas of excess local negative charge, and therefore candidates for conversion to Lys. In this 179-residue protein, the replacement of an anionic residue by Lys corresponds to a pI increase of about 0.3.

Five mutants were produced for this study: M18K, D19N, D19A, E25K, and D55K (Figure 4). The D19 mutants expressed poorly and mostly in the pellet fraction, likely due to some unforeseen, important folding or solubility effect of the Asp 19 sidechain, and were not purified. The other 3 were purified and subjected to crystal screening; all three crystallized, generating several new crystal forms. All but one of the resulting crystal forms grew too small and/or compound for diffraction, but about half of them crystallized above pH 6. This compares with

wild-type, which, after more extensive screening than the mutants, yielded only two forms, both below pH 5¹³. E25K crystallized only at pH values below and near 5, but the six observed M18K crystal forms (out of 150 conditions screened) grew at pH values of 4.5, 4.6, 7.5, 8.0, 8.5, and 8.5, while the eight observed D55K crystal forms grew at pH values of 4.6, 4.6, 5.6, 6.5, 6.5, 7.5, 7.5, and 7.5.

One of these D55K forms growing at pH 6.5 grew large and single enough for diffraction. A photo of this monoclinic form is shown in Figure 5. These crystals belong to space group $P2_1$ with unit cell 64.5 Å, 38.1 Å, 80.7 Å, 98.7°. As with the wild-type, these crystals tended to grow as compound plates, but their thickness was significantly better, leading to the 1.6 Å resolution structure 3ghx.pdb. This structure at the D55K mutation site shows the importance of the mutation to the growth of this crystal form: Figure 6 shows that one crystal contact features a salt link between the new lysine 55 and Asp 156. The region of Asp 156 also contains Glu 153, so that, without the mutation, both sides of this contact would have negative charges, making it unlikely that the contact could form near neutral pH in the wild-type protein.

4. Discussion

This work has demonstrated one case of a low-pI protein whose crystallization was limited to low pH, and the best single crystal form that was observed, was growth-limited (to thin plates) by a carboxyl-carboxyl hydrogen bond in the corresponding crystal contact. These limitations (crystallization pH profile and crystal morphology) presented serious obstacles to solving the structure and to further study of the enzyme's mechanism. The limitations were overcome by mutations that raised the pI. In the refined structure of the D55K mutant, the mutant Lys residue participates in a well-ordered crystal contact.

About 12 % of proteins have pI values below 5, and the relatively low crystallographic success rates²⁰ for proteins at both extremes of pI are likely related to their unbalanced surface charge. Proteins with low pI values have, under most conditions, net anionic surfaces whose excess negative charge disfavors formation of complementary contacts. While the gross protein-protein repulsion can be overcome by salt, acidic side chains may not form well-ordered contacts efficiently with each other because of the large number of conformational degrees of freedom in cation-bridged contacts. At low pH, acidic side chains can form carboxyl-carboxyl hydrogen bonds, conferring pH-dependence on growth and morphology. The observation of these dicarboxyl bonds in crystal contacts provides an atomic-level explanation for the limitation to low-pH growth conditions, and suggests that mutations to remove excess negative charge may increase the likelihood of forming favorable contacts⁶.

These considerations can potentially be applied to other proteins as follows. Nonconserved Asp and especially Glu residues that are predicted to occur at exposed turns and loops (e.g., tandem acidic residues) are the best targets. In cases of low-pI proteins that have failed to crystallize, if solubility is good, then Glu → Ala or → Ser mutants will provide significant surface entropy reduction as well as excess charge reduction. If low solubility is a problem, then converting the acidic residues to Lys will be more likely to prevent loss of solubility as well as more effective in balancing the surface charge. If a low-pI protein crystallizes only (or preferentially) at low pH, then excess negative charge is likely a factor and Lys mutants may provide the shortest path to balancing the excess charge. In principle, similar considerations (with opposite charges) should apply also for recalcitrant high-pI proteins; in these cases excess surface charge may be most effectively mitigated by K → D mutations, which would also reduce entropy.

For low-pI proteins, the strategy of excess charge reduction (ECR), represented by the mutations E → K and D → K, may be compared with surface entropy reduction (SER) and

with solubility-enhancing strategies. SER mutations tend to reduce solubility and thus highlight an apparent paradox: surface entropy has been shown to disfavor crystallization²¹, but the charged high-entropy side chains EK are also known to confer solubility^{22,23}, which is vital for crystallization. The optimal balance between surface entropy's beneficial effects (for solubility) and its detrimental effects on crystallization is difficult to establish. Along with surface entropy, excess net surface charge also appears to inhibit crystallogenesis, but the charge-flipping ECR mutations do not compromise solubility as much as SER mutations. Along with SER strategies for highly soluble proteins and solubility-enhancement for poorly soluble proteins, the ECR approach for low-pI cases is another tool available for enhancing crystallizability of recalcitrant proteins.

Acknowledgments

The authors gratefully acknowledge the support of the National Institutes of Health (NCRR) for data collected at NSLS beamline X29, and the technical and logistical assistance of Darwin Diaz and Dawn Rode of the Center for Advanced Research in Biotechnology. Disclaimer: identification of specific instruments and products in this paper is solely to describe the scientific procedures and does not imply recommendation or endorsement.

Abbreviations

AC	adenylyl cyclase
DTT	dithiothreitol
IPTG	isopropylthiogalactoside
MS	Mass Spectrometry
NCS	noncrystallographic symmetry
PAGE	polyacrylamide gel electrophoresis
PEG	polyethylene glycol
RT	room temperature

References

1. Derewenda ZS. *Methods* 2004;34:354–363. [PubMed: 15325653]
2. Derewenda ZS. *Structure* 2004;12:529–535. [PubMed: 15062076]
3. Dyda F, Hickman AB, Jenkins TM, Engelman A, Craigie R, Davies DR. *Science* 1994;266:1981–1986. [PubMed: 7801124]
4. Das D, Georgiadis MM. *Protein Sci* 2001;10:1936–1941. [PubMed: 11567084]
5. Davies DR, Hickman AB. *International Tables for Crystallography* 2006;F:100–110. ch. 4.3.
6. Gallagher DT, Pan Q, Gilliland GL. *J Crystal Growth* 1998;193:665–673.
7. Derewenda Z, Vekilov P. *Acta Cryst* 2006;D62:116–124.
8. Pantazatos D, Kim JS, Klock HE, Stevens RC, Wilson IA, Lesley SA, Woods VL Jr. *PNAS* 2004;101:751–756. [PubMed: 14715906]
9. Dasgupta S, Iyer GH, Bryant SH, Lawrence CE, Bell JA. *Proteins* 1997;28:494–514. [PubMed: 9261866]
10. Jenkins TM, Hickman AB, Dyda F, Ghirlando R, Davies DR, Craigie R. *PNAS* 1995;92:6057–6061. [PubMed: 7597080]
11. Dale GE, Broger C, Langen H, D'Arcy A, Stiiber D. *Protein Eng* 1994;7:933–939. [PubMed: 7971955]
12. Stover C, Mayhew MP, Holden MJ, Howard A, Gallagher DT. *J Struc Biol* 2000;129:96–99.
13. Smith N, Kim SK, Reddy PT, Gallagher DT. *Acta Cryst* 2006;F62:200–204.
14. Gallagher DT, Smith NN, Kim SK, Reddy PT. *J Mol Biol* 2006;362:114–122. [PubMed: 16905149]
15. McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ. *Acta Cryst* 2005;D61:458–464.

16. Murshudov GN, Vagin AA, Dodson EJ. *Acta Cryst* 1997;D53:240–255.
17. McRee D. *J Struct Biol* 1999;125:156–165. [PubMed: 10222271]
18. Hartmann, P. *Growth of Crystals*. Sheftal, NN., editor. Vol. 7. Consultants Bureau; New York: 1969. p. 3-18.
19. Strom CS, Bennema P. *J Crystal Growth* 1997;173:150–158.
20. Canaves JM, Page R, Wilson IA, Stevens RC. *J Mol Biol* 2004;344:977–991. [PubMed: 15544807]
21. Price WN II, Chen Y, Handelman SK, Neely H, Manor P, Karlin R, Nair R, Liu J, Baran M, Everett J, Tong SN, Forouhar F, Swaminathan SS, Acton T, Xiao R, Luft JR, Lauricella A, DeTitta GT, Rost B, Montelione G, Hunt JF. *Nature Biotech* 2009;27:51–57.
22. Smialowski P, Martin-Galiano AJ, Mikolajka A, Girschick T, Holak TA, Frishman D. *Bioinformatics* 2007;23:2536–2542. [PubMed: 17150993]
23. Trevino SR, Scholtz JM, Pace CN. *J Pharmaceutical Sci* 2008;97:4155–4166.

Figure 1. Wild-type triclinic crystals (cell 33.5, 35.5, 72.0 Å), typical clusters of thin plates, growing at pH 4.6. The largest single crystallite is about 100 microns wide and 5 microns thick. The large (001) face is growth-limiting, and the difficulty of growing thicker crystals (which, for this form, corresponds to stronger diffraction) was a major obstacle to solving the structure.

Figure 2.

Packing in the *ab* plane of the triclinic crystals. Molecules in this plane have a hexagonal-close-packed arrangement, such that each molecule has 6 neighbors in the plane. The dimers in this view are seen end-on, so that behind each visible A-chain is a hidden B-chain (related by noncrystallographic symmetry). Thus the long axis of the dimer is along the crystal's *c* direction. Asp A19 in the central molecule is shown projecting outward (toward the viewer). Each Asp A19 forms a hydrogen bond to Asp B19 in the next higher layer of the crystal, by the contact shown in Figure 3.

Figure 3. Stereo view showing the dicarboxyl hydrogen bond in the c-directed crystal contact in the wild-type structure 2FJT. This type of hydrogen bond can occur only at low pH, when one of the two carboxylates is protonated. The 2Fo-Fc electron density map is shown contoured at 1.1 sigma.

Figure 4. Ribbon representation of the dimer showing the locations of the mutations. All were observed to be on exterior convex regions of the protein surface in the wild-type structure.

Figure 5.

Monoclinic crystal of the D55K mutant growing at pH 6.5. Width of the field is about 200 μm . Although this crystal has many surface satellites, it has a thickness of over 25 μm and consists mostly of a single unit. Similar and slightly better (with larger single domains) crystals were soaked with substrate analogs to generate active-site complexes diffracting to about 1.6 $\text{\AA}$.

Figure 6.

Stereo view showing the hydrogen bond from the mutated Lys 55 residue (left) to Asp 156 on a neighboring molecule. This bond could not have occurred in wild-type. The molecule on the right has two carboxylates in the region (Asp 156 and Glu 153), hence this crystal contact would be unlikely to occur at neutral pH without the mutation. 2Fo-Fc electron density is shown contoured at 1.3 sigma.

Figure 7.