The crystal structure of CREG, a secreted glycoprotein involved in cellular growth and differentiation

Michael Sacher*†‡, Alessandra Di Bacco§, Vladimir V. Lunin¶‖, Zheng Ye*†, John Wagner*†, Grace Gill‡§, and Miroslaw Cygler*†¶

*Montreal Proteomics Network, 740 Doctor Penfield, Montreal, QC, Canada H3A 1A4; †Department of Biochemistry, McGill University, 3655 Promenade Sir William Osler, Montreal, QC, Canada H3G 1Y6; §Department of Pathology, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115; and ¶Biotechnology Research Institute, 6100 Royalmount Avenue, Montreal, QC, Canada H4P 2R2

Edited by Brian W. Matthews, University of Oregon, Eugene, OR, and approved November 3, 2005 (received for review June 17, 2005)

The cellular repressor of E1A-stimulated genes (CREG) is a secreted glycoprotein that inhibits proliferation and enhances differentiation of human embryonal carcinoma cells. CREG binds to the cation-independent mannose 6-phosphate (M6P)/insulin-like growth factor II (IGF2) receptor (IGF2R) (M6P/IGF2R), and this receptor has been shown to be required for CREG-induced growth suppression. To better understand CREG function in cellular growth and differentiation, we solved the 3D crystal structure of this protein to 1.9-Å resolution. CREG forms a tight homodimeric complex, and CREG monomers display a β-barrel fold. The three potential glycosylation sites on CREG map to a confined patch opposite the dimer interface. Thus, dimerization of glycosylated CREG likely presents a bivalent ligand for the M6P/IGF2R. Closely related structural homologs of CREG are FMN-binding split-barrel fold proteins that bind flavin mononucleotide. Our structure shows that the putative flavin mononucleotide-binding pocket in CREG is sterically blocked by a loop and several key bulky residues. A mutant of CREG lacking a part of this loop maintained overall structure and dimerization, as well as M6P/IGF2R binding, but lost the growth suppression activity of WT CREG. Thus, analysis of a structure-based mutant of CREG revealed that binding to M6P/IGF2R, while necessary, is not sufficient for CREG-induced growth suppression. These findings indicate that CREG utilizes a known fold.

cell growth | cell proliferation | cellular repressor of E1A-stimulated genes

The human teratocarcinoma cell line NTERA-2 is a useful model system to study tumorigenesis and early embryonic development, because it shares common features with pluripotent embryonic cells (1). This is underscored by the fact that this cell line can be induced to differentiate into neuronal and nonneuronal phenotypes by many different compounds (2). Treatment of NTERA-2 cells with hexamethylene bisacetamide (HMBA) induces differentiation into a nonneuronal phenotype (2), whereas treatment with retinoic acid induces a neuronal phenotype (2, 3). Although the pathways involved in these differentiation events are yet to be fully elucidated, it is interesting that these two compounds appear to activate common pathways, as suggested by both up- and down-regulation of common downstream targets (4, 5).

One common protein up-regulated by these differentiation-inducing reagents is the cellular repressor of E1A-stimulated genes (CREG) (6, 7), a secreted glycoprotein of 220 amino acids (human isoform). CREG transcript levels increase during embryonic development, and the protein is progressively secreted into the medium during HMBA-mediated *in vitro* differentiation of NTERA-2 cells as a 191-residue-long mature form (6). Overexpression of CREG enhances both HMBA- and retinoic acid-mediated NTERA-2 differentiation and, when applied ectopically in the growth medium, can enhance spontaneous cellular differentiation. Consistent with its antiproliferative effect, CREG overexpression increases the percentage of cells in G_1 phase in an unsynchronized population of NTERA-2 cells and leads to fewer and smaller colonies, as judged by growth in soft agar (8). It has been proposed, therefore, that CREG acts as a ligand that enhances differentiation and/or reduces cell proliferation. Interestingly, in cardiomyocytes, CREG appears to regulate the levels of the signaling kinases ERK1/2 (9), suggesting the potential involvement of the mitogen-activated protein kinase pathway in the mechanism of CREG-induced changes in cell growth and differentiation.

Of the several proteins that have been reported to interact with CREG, the cation-independent mannose six-phosphate (M6P)/insulin-like growth factor II (IGF2) receptor (IGF2R) (M6P/IGF2R) has been shown to be required for the growth-suppressive activity of CREG (7, 8, 10). Endogenous CREG localizes to discrete punctate structures, which do not colocalize with the lysosomal protein LAMP-1 (8), within the cell in an M6P/IGF2R-dependent fashion. Several lines of evidence support a role for the M6P/IGF2R in the regulation of cell growth and tumorigenesis. First, by binding to and internalizing IGF2, the receptor down-regulates the mitogenic effects of this growth factor (11). Second, the receptor participates in the activation of the growth suppressor TGF-β (12, 13). Third, the receptor has been found to be mutated in many different forms of cancer, including breast, liver, lung, and ovarian cancers (see ref. 14 and refs. therein).

Understanding how CREG and M6P/IGF2R cooperate in promoting cellular differentiation and arrest of proliferative cell growth will shed light on the mechanisms of carcinogenesis and embryonic development. As a first step in this endeavor, we have solved the 3D crystal structure of CREG. The protein associates into tightly bound dimers, with each monomer forming a β-barrel structure similar to the FMN-binding split-barrel fold observed in some oxidases and oxidoreductases. We have found that the putative FMN-binding pocket in CREG is occupied by a loop region, which, while apparently blocking FMN binding, is functionally important for the ability of CREG to suppress cell growth.

Materials and Methods

Cloning of CREG. The CREG construct used for structure determination included residues 49–220 of the human protein. The resulting expression plasmid, designated pZY50, has an ORF that includes CREG residues 49–220 and an N-terminal leading sequence of MGSSHHHHHHGS.

Expression of Recombinant CREG. Bacterial BL21(DE3) cells containing the plasmid pZY50 were used to inoculate LB medium

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: CREG, cellular repressor of E1A-stimulated genes; IGF2, insulin-like growth factor II; IGF2R, IGF2 receptor; M6P, mannose-6-phosphate; TCA, trichloroacetic acid; PDB, Protein Data Bank; SeMet, selenomethionine.

Data deposition: The coordinates for CREG have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 1XHN).

‡To whom correspondence may be addressed. E-mail: michael.sacher@bri.nrc.ca or grace_gill@hms.harvard.edu.

‖Present address: Banting and Best Department of Medical Research, University Health Network, 112 College Street, Toronto, ON, Canada M5G 1L6.

www.pnas.org/cgi/doi/10.1073/pnas.0505071102

containing 50 μg/ml ampicillin and grown for ≈2 h at 37°C. Protein production was induced at an OD_{600} of ≈1.0 by the addition of isopropyl-β-D-thiogalactopyranoside to 200 μM, and the culture was allowed to grow for an additional 20 h at 20°C. Cells were harvested and the cell pellet was stored at −20°C. Lemaster-defined media containing 25μg/ml selenomethionine (SeMet) was used for SeMet derivative protein production (15) in the methionine auxotrophic DL41 *E. coli* containing the plasmid pZY50. The culture was grown as above.

Protein Purification. Frozen cells were thawed and resuspended in lysis buffer (50 mM Tris, pH 7.5/400 mM NaCl/20 mM imidazole/5 mM 2-mercaptoethanol/5% glycerol/1% Triton X-100/one tablet of protease inhibitor mixture without EDTA (Roche)/30 ml of buffer) followed by sonication on ice. The lysate was centrifuged at 45,000 rpm, the supernatant was combined with 2 ml of a nickel (II)–nitrilotriacetic acid (Ni–NTA) agarose resin equilibrated in lysis buffer, and binding took place on a rotating wheel at room temperature, after which the entire content was poured into a PolyPrep Bio-Rad column. The resin was washed successively with wash buffer I (lysis buffer without protease inhibitor and with 50 mM Tris, pH 8.0), wash buffer II (wash buffer I with 1 M NaCl and 40 mM imidazole), and wash buffer III (wash buffer II without Triton X-100 and with 400 mM NaCl). Proteins were eluted with elution buffer (200 mM imidazole/50 mM Tris, pH 8.0/200 mM NaCl/5 mM 2-mercaptoethanol/5% glycerol) for 10 min. Elution was repeated three more times, and the fractions were pooled. The eluted protein was concentrated, and the buffer was exchanged for 150 mM NaCl/50 mM Tris, pH 7.5/1 mM 2-mercaptoethanol. The final protein solution was concentrated to 8.7 mg/ml. The homogeneity of the protein solution was tested by dynamic light scattering on a Proterion DynaPro MSPR II plate reader (Proterion, Piscataway, NJ).

The SeMet-labeled protein was purified by the same procedure. The level of SeMet incorporation into the protein was determined by using a liquid chromatography/electrospray ionization mass spectrometer (1100 Series, Agilent Technologies, Palo Alto, CA).

Crystallization. Initial trials were performed by the hanging-drop vapor diffusion method with commercially available screens purchased from Hampton Research (Riverside, CA). Optimization trials were performed by the hanging-drop method with optimal conditions for crystal growth containing 15% polyethylene glycol (PEG) 4000/0.1 M sodium acetate (NaOAc), pH 5.0. To obtain larger and single crystals, the crystals were microseeded into a solution of 16% PEG 4000/0.1 M NaOAc, pH 5.0/15–20% ethylene glycol. These crystals belong to the space group $P2_12_12_1$ with cell dimensions a = 56.1, b = 106.2, and c = 120.9 Å and contain four molecules in the asymmetric unit. The purified SeMet-labeled protein (9.0 mg/ml) was crystallized as described for the native protein and gave crystals with the same cell dimensions. Before data collection, the crystal was briefly immersed in a cryoprotecting solution containing 25% PEG 4000, 12.5–15% ethylene glycol/0.1 M NaOAc, pH 5.0, immediately flash-frozen in a stream of N_2 gas (100 K) on a goniometer head, and used for data collection.

Data Collection and Refinement. Diffraction data were collected at the X8C synchrotron beamline at Brookhaven National Laboratory, Upton, NY, by using a Quantum Biotechnologies-4 (Montreal) charge-coupled device detector. Three-wavelength multiwavelength anomalous datasets about the Se K edge resolution were collected from Se-Met-labeled apoprotein to a resolution of 1.95 Å. The data were processed with the HKL2000 program suite (16).

The structure of CREG was determined by using datasets collected at peak, inflection point, and remote wavelengths. All 16 expected Se sites were found by using the program SOLVE (17). Density modification and model building were performed by using RESOLVE (18), resulting in an increase of the figure of merit from 0.49 to 0.69. The model assembled by the RESOLVE procedure was almost complete and contained >80% of residues. Further model building and refitting were performed by using the program O (19), alternating with cycles of refinement using the program REFMAC5 (20). The model was refined to a final R factor of 0.163 and R_{free} of 0.214 at 1.95-Å resolution. The polyhistidine tag and the N-terminal 10–11 residues of the recombinant construct were disordered in the crystal and were not modeled. Similarly, two C-terminal residues were not observed in the electron density. The final model contains

Table I. Data collection and refinement statistics

Measurement	Inflection	Peak	Remote
Data collection			
Wavelength, Å	0.980005	0.979502	0.972318
Cell: a, Å	106.15	106.13	105.98
b, Å	121.28	121.27	121.08
c, Å	55.92	55.91	55.835
Resolution (last shell), Å	50–1.95 (2.02–1.95)	50–1.95 (2.02–1.95)	50–1.95 (2.02–1.95)
R_{sym} (last shell)	0.095 (0.504)	0.091 (0.459)	0.093 (0.489)
Completeness (last shell), %	100 (100)	100 (100)	100 (100)
$I/\sigma(I)$ (last shell)	9.2 (2.8)	9.9 (3.5)	9.4 (3.0)
No. of reflections	549,899	549,280	550,973
Unique reflections	101,790	101,762	101,256
Refinement statistics			
R/R_{free}		0.163/0.215	
No. of non-H protein atoms, molecule A, B, C, D		1,371, 1,362, 1,378, 1,399	
No. of water molecules		896	
B factor, Å^2: chain A, B, C, D			
Main-chain atoms		15.6, 18.5, 18.5, 18.9	
Side-chain atoms		17.7, 20.6, 20.3, 20.8	
Water molecules		34.1	
rmsd bond length, Å		0.017	
rmsd bond angle, degrees		1.478	
Ramachandran plot, residues			
In most-favorable region, %		91.8	
In disallowed regions, %		0.0	

rmsd, rms deviation.

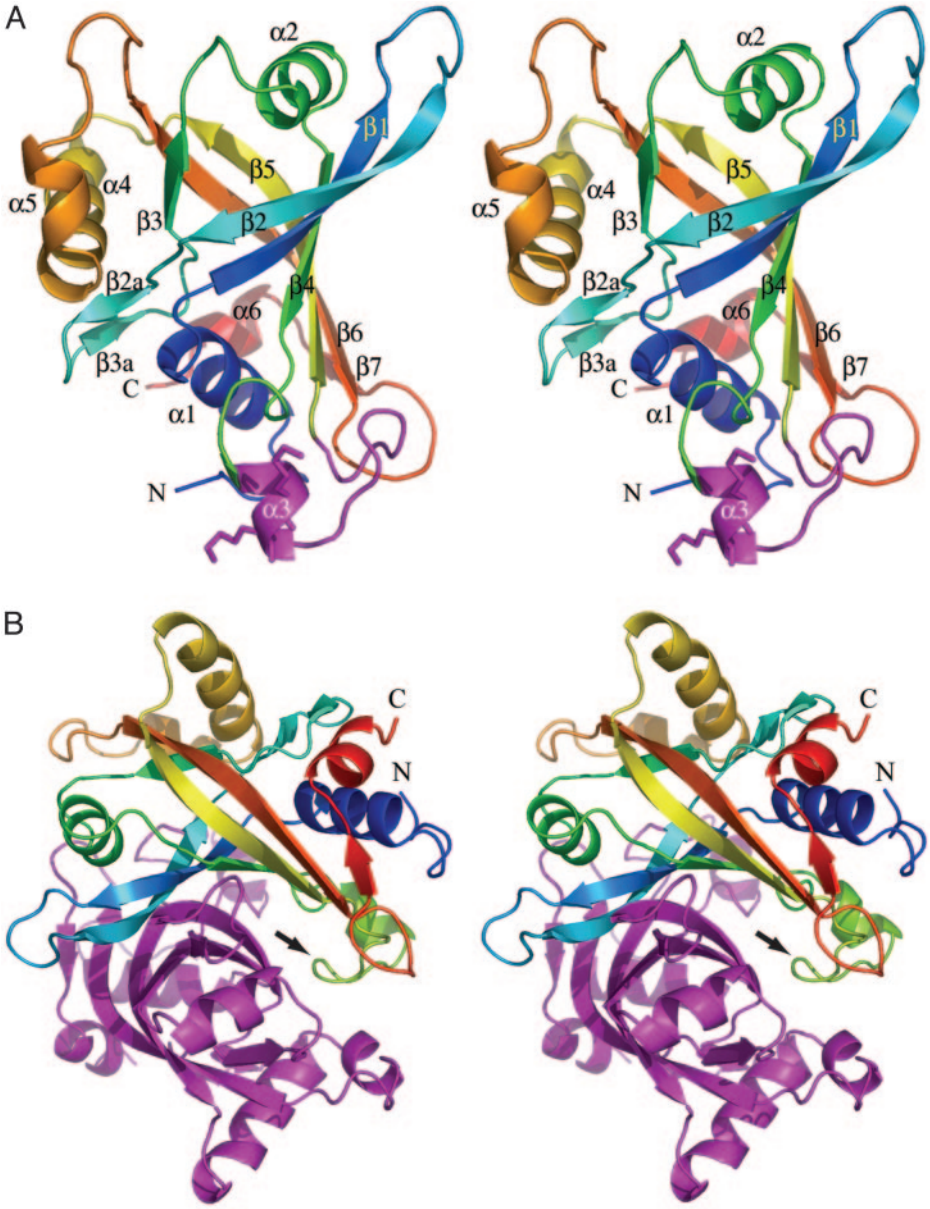

Fig. 1. The structure of CREG. (*A*) A stereoview of the cartoon representation of the CREG monomer displaying the β-barrel fold. The ribbon is colored by secondary structure progression from blue at the N terminus to red at the C terminus. Secondary structural elements as well as the N and C termini are indicated. Helix α3, with lysines 136 and 137, and the subsequent loop are colored in magenta. (*B*) A stereoview of the cartoon representation of the CREG dimer. One molecule is colored by secondary structure progression from blue at the N terminus to red at the C terminus, and the second molecule is colored in magenta. The loop region in green (indicated by an arrow) occluding the FMN pocket of the magenta-colored molecule is clearly visible. This figure was rendered with the program PYMOL.

four molecules in the asymmetric unit and includes residues 60–218 (molecules A and D) or 61–218 (molecules B and C) (numbering according to SwissProt entry O75629) and 892 water molecules. The model has good geometry without outliers, as shown by the program PROCHECK (21). Data collection statistics are shown in Table 1.

Assays of CREG Expression and Function in Mammalian Cells. WT and mutant forms of CREG expressed from the cytomegalovirus (CMV) promoter were transfected into NTERA-2 cells, and colony formation was assayed as described (8). One-way ANOVA and the Bonferroni method for calculation of the least-significant difference were considered to be statistically significant. The two CREG mutants were (*i*) K136A/K137A (CREG M1) and (*ii*) Δ141–144 (CREG M2). 293T cells were transfected with 6 μg of CMV, WT CREG, CREG M1, or CREG M2 together with GFP. SDS total cell lysates were harvested 48 h after transfection, and media were trichloroacetic acid (TCA)-precipitated. CREG levels were measured with a polyclonal antibody raised against His-6-tagged human CREG (1:500). Anti-GFP (1:400) was used as a transfection marker.

Binding of IGF2R to CREG WT and mutants was assayed by using Far Western blotting analysis, essentially as described (8) with 0.5 μg/ml purified soluble M6P/IGF2R (kindly provided by R. G. MacDonald, University of Nebraska, Lincoln) and anti-M6P/IGF2R antibody 1:2,000 (kindly provided by P. Lobel, Robert Wood Johnson Medical School, New York). Horseradish peroxidase-conjugated anti-rabbit antibody 1:1,000 (Promega) and enhanced chemiluminescence (PerkinElmer Life Science) were used for detection.

Results and Discussion

Overall Structure of CREG Monomer. The structure of CREG was solved by multiwavelength anomalous dispersion of a SeMet-substituted crystal. Of the 172 residues, 11 N-terminal residues and two at the C terminus were disordered and not visible in the electron-density map. The protein structure belongs to an α/β fold. Six long and twisted strands form an antiparallel β-barrel structure

constricted in the center and opened at both ends (Fig. 1*A*). The strands are interspersed with three two- to three-turn (α1, α2, and α4) and three short one-turn α-helices. An additional short two-stranded β-sheet is inserted in the connection between strands β2 and β3, and one short strand is added at the C terminus to one end of the barrel. The order of secondary structural elements is α1-β1 ↑ -β2 ↓ -β2a ↓ -β3a ↑ -β3 ↑ -α2-β4 ↓ -α3-β5 ↑ -α4-α5-β6 ↓ -β7 ↑ -α6. Helices α1 and α2 cap the two ends of the barrel. The interior of the barrel is filled with hydrophobic residues, and there is an extensive hydrogen-bond network between the β-strand mainchain atoms. In addition, several salt bridges provide additional stabilization to the structure. In particular, Asp-67. . . Arg-174. . . Asp-93. . . His-64 ties the C-terminal end of helix α4 to strands β1, β2, and β2a; Glu-78. . . Arg-81 stabilizes the connection between β2a and β3a; and Arg-60 forms hydrogen bonds to several backbone carbonyl groups fixing helix α1 in the capping position. This tightly packed hydrophobic core and numerous hydrogen bonds provide substantial stability to the overall structure of the protein. This is evidenced by very small changes in the circular dichroism spectra, and as a consequence in the secondary structure of the protein, with urea concentrations of up to 6 M. The monomer has a smooth concave surface on one side, whereas the rest of the surface is ragged. This concave surface is the site of dimerization (see below). A disulfide bond is formed between Cys-135 and -147. This serves to orient the loop between helix α3 and strand β4 and restrict its flexibility. A portion of this loop was shown to contribute to the function of CREG in cell growth inhibition as described below.

Dimer Formation. The monomers of CREG associate into dimers (Fig. 1*B*). The dimerization is observed not only in the crystal but also in solution (see Fig. 4*A*). The dimer interface has largely hydrophobic character, with alanines, prolines, phenylalanines, and tryptophanes being at the center of these interactions. Several polar residues are also present, including threonines and to a lesser extent serines and glutamines. There are also 12 ordered water molecules in the interface providing networking hydrogen bonds to the main- and side-chain atoms of both monomers. The dimer-forming interactions extend over a rather large surface area of ≈1,800 Å^2 in each monomer, which corresponds to 20% of the ≈9,100-Å^2 total surface area of the CREG molecule.

Sequence and Structural Similarities to Other Proteins. To gain insight into the function of CREG, detailed sequence and structure-based comparisons were performed. Two CREG isoforms have been identified so far in humans, which show 39% sequence identity in the region of the molecule corresponding to the determined 3D structure (7, 10). Given the level of identity between this region of the proteins and CREG, we predict that these related proteins will be structurally similar to CREG. A search through sequence databases shows the presence of CREG homologs not only in mammals but also in metazoa such as *Drosophila melanogaster* and *Anopheles gambiae* and in plants (*Arabidopsis thaliana*). A CREG homolog sharing 46% sequence identity was also identified in zebrafish. These proteins vary in length from ≈170 up to ≈600 residues and share a region of ≈160 residues. This common region is predominantly located in the C terminus of these proteins, except in plants where it forms the middle domain. A comprehensive PSI-BLAST search (22) identified many other predominantly bacterial protein sequences with a sequence identity in the 12–18% range and with a spectrum of functional annotations such as heme iron utilization protein, pyridoxamine 5′-phosphate oxidase (FMN flavoprotein, by similarity), 5-nitroimidazole antibiotic resistance protein (FMN flavoprotein), and flavin-nucleotide-binding protein. Because many of these sequences have been derived from translation of genomic DNA and annotated based on sequence similarity, it is not known whether all of them indeed bind flavin cofactors. Furthermore, we find no common functional denominator for these proteins. Thus, this analysis indicates that sequence identity alone does not shed light on the function of CREG.

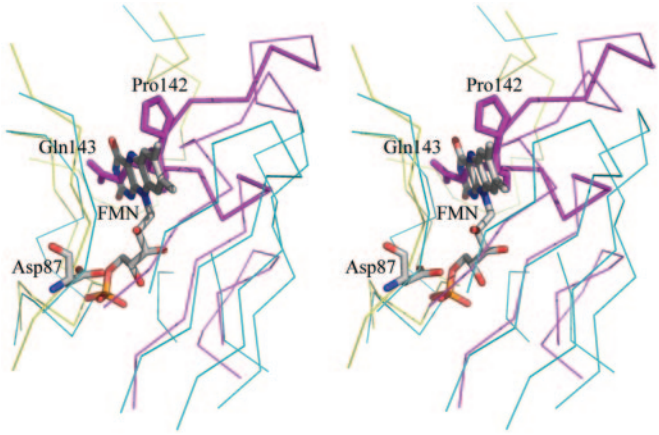

Fig. 2. The blocked FMN pocket on CREG. A ribbon representation of the superposition of CREG with the yeast pyridoxine-5′-phosphate oxidase (PDB ID code 1CI0) showing the loop that blocks the FMN pocket (shown in thicker lines). Only residues within 5 Å of the FMN molecule (painted in atom colors) are shown. The CREG dimer is painted yellow (chain A) and magenta (chain B), and pyridoxine-5′-phosphate oxidase is painted in cyan. Two side chains, Pro-142 and Gln-143 from chain B, protruding into the FMN space are shown in full. Asp-87 from chain A is proximal to the position of the FMN phosphate. The alignment was performed with SWISS PDB VIEWER, and the figure rendered with PYMOL.

A search for proteins with similar 3D structures conducted with DALI (23) revealed several proteins with a similar fold as CREG. Their fold is classified as an FMN-binding split barrel in the SCOP database (24). The closest structural homologs are proteins of bacterial origin that include 5-nitroimidazole antibiotic resistance protein NimA from *Deinococcus radiodurans* [Protein Data Bank (PDB) ID code 1W3O], pyridoxine 5′-phosphate (PNP) oxidase from *Saccharomyces cerevisiae* (PDB ID code 1CI0), FMN-binding protein from *Desulfovibrio vulgaris* (PDB ID code 1AXJ), human protein with predicted PNP oxidase activity (PDB ID code 1NRG), PNP oxidase from *Pseudomonas aeruginosa* (PDB ID code 1T9M), and several other FMN-binding oxidases and oxidoreductases as well as other proteins with unknown functions. All these proteins form dimers, and the FMN cofactor-binding site was localized to the interface between the two monomers.

Structural similarity to proteins known to bind flavin–nucleotide cofactor prompted us to test whether CREG indeed binds FMN. We obtained crystals of CREG from solution containing 1 mM FMN and in addition incubated native crystals in solution containing 1 mM FMN cofactor. Data collected from these crystals showed no presence of FMN bound to the protein. We also attempted to identify spectroscopically the possible binding of FMN. To this end, we incubated recombinant CREG with increasing concentrations of FMN (1–10 mM), removed the excess FMN by desalting the solution, and examined the resulting protein by spectroscopy between 300- and 700-nm wavelength. As a positive control, we used the *E. coli* protein PAD1 from bacteria that has been shown to bind FMN *in vitro* (25). FMN gives characteristic peak absorbances at ≈380 and 450 nm. Although binding of FMN to PAD1 was readily detected by this methodology, these experiments showed no evidence for CREG binding the FMN cofactor (data not shown). Comparison of the structure of CREG with pyridoxine 5′-phosphate oxidases complexed with FMN showed that CREG has a ≈15-residue insertion in the loop between strands β4 and β5 that protrudes into the space occupied by FMN (Fig. 2). Furthermore, two arginines involved in binding the FMN-terminal phosphate group are replaced by a methionine and an aspartate, respectively,

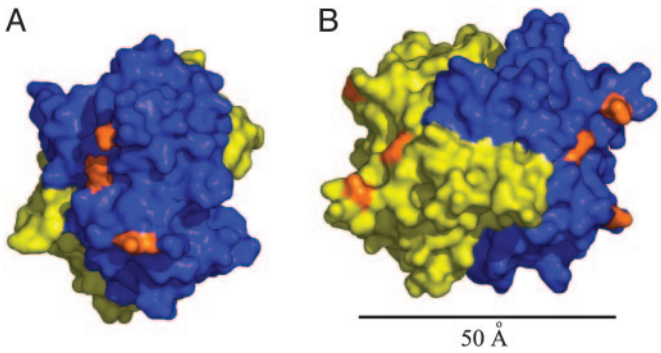

Fig. 3. The glycosylation sites on CREG. (*A*) A view of the three potential *N*-glycosylation sites on one face of the CREG dimer. One molecule is colored in blue, and the second molecule is colored in yellow. The three potentially glycosylated Asn residues (160, 193, and 216) are colored in orange. (*B*) A relative disposition of the two *N*-glycosylation patches on opposite faces of the CREG dimer. The molecule in *A* was rotated ≈90° around the *y* axis. The distance between the nearest asparagine residues is ≈50 Å. *B* was rendered with PYMOL and is colored as in *A*.

suggesting that this site even in the absence of the β4–β5 extended loop would not likely accommodate the phosphate.

Therefore, we conclude that CREG, although having a fold associated with FMN binding, does not bind this cofactor. Indeed, the recently determined structure of a protein with unknown function from *Anabaena* sp. PCC 7120 (PDB ID code 1VL7), which is the closest CREG structural homolog, also has a longer loop between strands β4 and β5 and, based on the superposition with FMN-binding proteins, most likely does not bind this cofactor. These results suggest that this fold, designated in the SCOP database (24) as FMN-binding split-barrel fold, has indeed been adapted in eukaryotes during evolution to perform other functions besides FMN-dependent oxidation reactions. As we discuss below, the functionally important area for the new cellular role acquired by CREG remains in the same part of the barrel fold that is used to bind the FMN cofactor in other proteins and is here filled by a long insertion forming the extended loop protruding into the putative FMN-binding pocket.

Structural Implications for Interaction with the M6P/IGF2 Receptor. A previous study has shown that CREG binds to the cation-independent M6P/IGF2R, and studies in cells lacking M6P/IGF2R expression revealed that this receptor is required for CREG-induced inhibition of cell growth (8). Furthermore, deglycosylation of CREG reduced its affinity for M6P/IGF2R, suggesting that the carbohydrate groups mediate this interaction. The three potential *N*-glycosylation sites on CREG (Asn-160, -193, and -216) all map to the surface of the protein opposite the dimerization interface (Fig. 3*A*) and are proximal to each other. Their glycosylation would have the effect of producing a M6P patch on a fairly restricted region of the molecule. In the context of the CREG dimer, these glycosylation patches are on the opposite ends of the elongated dimer molecule at a distance of >50 Å from each other (Fig. 3*B*). Thus, although monomeric glycosylated CREG may be sufficient to bind the M6P/IGF2R (ref. 8 and Fig. 4), models derived from our structure raise the possibility that dimerization of glycosylated CREG may increase affinity and/or clustering of the CREG–receptor complex through interactions with two distinct M6P-binding pockets. It is noteworthy that M6P/IGF2R contains three M6P-binding pockets, two high-affinity sites found in repeat domains 1–3 and 9 and a low-affinity site in repeat domain 5 (26–28). Although the proximity of these M6P pockets in the M6P/IGF2R 3D structure is not known, modeling suggests that the high-affinity sites can come to within ≈45 Å of each other (29). Because the diameter of the CREG dimer is of a similar size, CREG could potentially interact simultaneously with both of these pockets. Another possibility that should be considered is a role of CREG in receptor dimerization. It has been shown that the receptor tends to dimerize (30), and that dimerization is mediated by multivalent, M6P-containing ligands (31). Furthermore, intermolecular crosslinking of the receptor by a multivalent protein

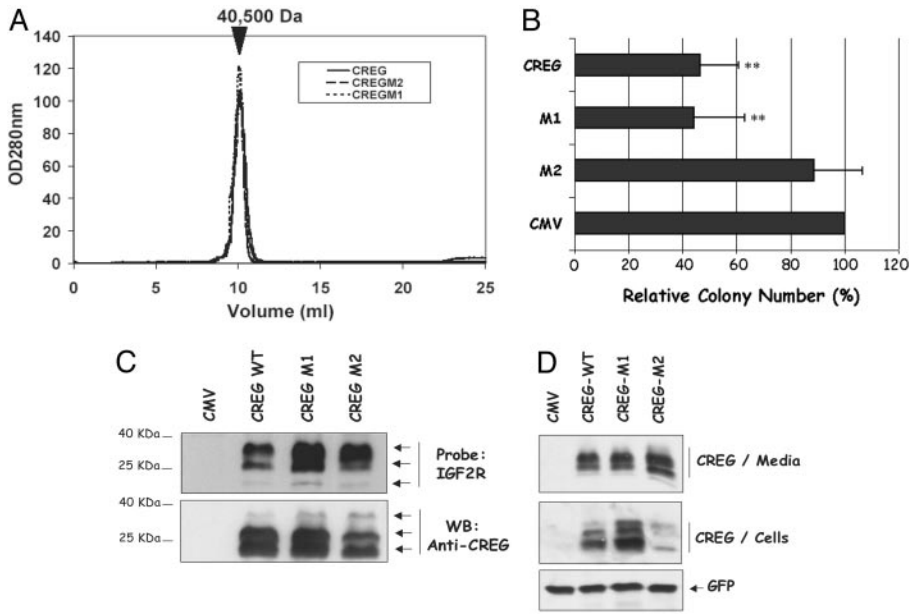

Fig. 4. Deletion of CREG residues 141–144 impairs activity. (*A*) Size-exclusion chromatography performed with WT and mutant CREG (M1, K136A/K137A and M2, Δ141–144) shows that the structural integrity of CREG and its ability to form dimers were not affected by the mutations. The approximate size of 40,500 Da is based on the migration of standards in the Superdex 75 (GE Healthcare) column used. (*B*) NTERA-2 cells were transfected with WT CREG, CREG M1, CREG M2, or vector alone, and their ability to form colonies was analyzed as described in *Materials and Methods*. A graphic representation of the growth inhibition induced by CREG is shown. Values are relative to cytomegalovirus (CMV) vector control, which was set at 100%. Bar graphs represent the mean percentage ± SEM of three independent experiments. CREG WT and CREG M1 have a statistically significant reduction of colony number, as compared with CMV (**, $P \leq 0.005$ ANOVA). CREG M2 was significantly different from CREG WT or M1 ($P \leq 0.005$ ANOVA) but not statistically different from vector control. (*C*) CREG WT and mutants bind to M6P/IGF2R. Media from cells transfected with vector, CREG WT, M1, or M2 expression plasmids were TCA-precipitated, separated by gel electrophoresis, and transferred onto nitrocellulose. The membrane was probed with purified soluble M6P/IGF2R followed by detection with anti-M6P/IGF2R antibody. Subsequently, the membrane was stripped and reprobed with anti-CREG antibody. M6P/IGF2R binds similarly to CREG WT, M1, and M2, with a preference for more highly glycosylated forms of CREG. (*D*) 293T cells were transfected with WT CREG, CREG M1, CREG M2, or vector alone, and intracellular and extracellular CREG proteins were detected by immunoblot by using an anti-CREG antibody. Secreted CREG proteins were analyzed after TCA precipitation from culture media. Intracellular CREG protein levels were analyzed from whole-cell lysates. Detection of cotransfected GFP was used as a transfection and loading control. Multiple forms of CREG derive from *N*-linked glycosylation (6).

ligand, and not intramolecular crosslinking, appears to be important for increased receptor uptake (31). We speculate that, by virtue of multivalent M6P clusters exposed on opposite sides of the dimer, CREG could induce dimerization of M6P/IGF2R and thus promote uptake of the receptor from the cell surface. Recent studies have suggested that CREG may regulate M6P/IGF2R levels in ileal epithelial cells (32). How the internalization of the CREG-M6P/IGF2R complex affects cell growth remains unclear.

Functional Consequences of Helix α3 Mutations. To begin to address the mechanism by which CREG inhibits cell growth, it is necessary to identify regions on the protein involved in this process. We focused our attention on helix α3 and the extended loop region following this helix for several reasons (see Fig. 1*A*). First, there are several well conserved residues in this region of the molecule including Cys-135, Lys/Arg-137, Asp-141, Pro-142, Pro-145, and Cys-147. Second, this helix projects away from the bulk of the protein and is well suited to be involved in a function of the protein mediated by a protein–protein interaction. Third, the loop following helix α3 covers the region where FMN binds on the other structurally related proteins (see above), and we reasoned that its removal might facilitate FMN binding. Finally, the loop is a sufficient distance from the potential glycosylation sites and, thus, should not interfere with CREG glycosylation.

Two CREG mutations were generated. The first (CREG M1) contains a double point mutation converting the α3 residues Lys-136 and -137 to alanine. These lysine residues are exposed on the surface of the protein. The second mutant (CREG M2) contains a deletion of residues 141–144, inclusive, thus removing part of the region that occludes the putative FMN-binding site. Size-exclusion chromatography (Fig. 4*A*) and dynamic light scattering (not shown) showed there was no difference in the behavior of the mutant proteins when compared with WT CREG, suggesting that the mutations do not interfere with overall structural integrity or dimer formation. Furthermore, *in vitro* spectroscopic analysis confirmed that neither the M1 double mutant nor the M2 deletion mutant display any detectable FMN binding (data not shown).

We then tested the ability of the CREG mutants to inhibit growth of NTERA-2 cells (8). As seen in Fig. 4, CREG M1 was as effective as WT CREG in suppressing cell growth as measured by colony number (Fig. 4*B*) and colony size (not shown), indicating that this pair of lysines is not necessary for the growth inhibitory properties of CREG. In contrast, CREG M2 failed to inhibit cell growth and was indistinguishable from the vector-alone control (Fig. 4*B*). This result indicates that a portion of the loop region, but not Lys-136 or -137, is required for the growth inhibitory function of CREG. It is interesting to note that only two residues within the deleted region are conserved (Asp-141 and Pro-142), suggesting that one or both of these residues is involved in CREG-mediated growth inhibition.

Growth inhibition by CREG depends on M6P/IGF2R (8). We therefore examined M6P/IGF2R binding by the CREG mutants. As shown in Fig. 4*C*, in a Far Western assay, secreted CREG WT, M1, and M2 all bound similarly to the M6P/IGF2R receptor. Consistent with previous findings (8), the M6P/IGF2R bound more efficiently to highly glycosylated forms of CREG. This finding is consistent with our structural model (Fig. 3), which places the potential glycosylation sites and loop in different regions of the monomer (Fig. 3*A*). Importantly, these data show that binding to the M6P/IGF2R is not sufficient for CREG-mediated growth suppression. Thus, it is possible that residues deleted in M2 are required for binding a second CREG receptor, or these residues may alter the functional properties or trafficking of the M6P/CREG/IGF2R complex necessary for growth inhibition.

In this regard, it is interesting to note that a recent study suggests that secreted CREG down-regulates M6P/IGF2R levels (32), and M6P/IGF2R has been shown to regulate the intracellular distribution of CREG (8). In immunoblot analysis, we have failed to observe any significant reduction of steady-state M6P/IGF2R levels upon overexpression of CREG WT, M1, or M2 (data not shown). Although immunoblot analysis of CREG WT, M1, and M2 in the media revealed no differences in the secretion and glycosylation of these CREG derivatives, we observed a reduction in levels of intracellular CREG M2 compared with WT (Fig. 4*D*). Because the amount of secreted proteins was comparable, this is unlikely to be due to a difference in expression levels but rather might reflect a difference in the ability of CREG M2 to be internalized through binding to M6P/IGF2R or other receptors.

Conclusion

Taken together, our analysis of CREG M2 indicates that this mutant has maintained its overall structure but is defective in CREG-mediated growth suppression. Our data suggest that CREG binds to M6P/IGF2R as a bivalent M6P-containing ligand, but CREG residues between positions 141 and 144 are required, in addition to M6P/IGF2R binding, for proper CREG trafficking and growth inhibition. Future studies to determine the biochemical defect of the CREG M2 mutant will provide additional insights into the mechanisms by which CREG functions to inhibit cell proliferation. Furthermore, the structural data suggest that eukaryotic cells can elegantly modify a protein fold used in one process to change the function of the resulting protein.

We are grateful to Christine Gadoury for technical assistance. The generosity of Richard G. MacDonald, who provided purified soluble IGF2R, and Peter Lobel, who provided IGF2R antisera, is gratefully acknowledged. This work was supported by funding through the Montreal Proteomics Network, Genome Canada, Genome Quebec, the National Research Council of Canada (to M.C.), and National Institutes of Health (to G.G.).

1. Andrews, P. W. (1998) *APMIS* **106,** 158–167.
2. Andrews, P. W., Gonczol, E., Plotkin, S. A., Dignazio, M. & Oosterhuis, J. W. (1986) *Differentiation (Berlin)* **31,** 119–126.
3. Lee, V. M. & Andrews, P. W. (1986) *J. Neurosci.* **6,** 514–521.
4. Wakeman, J. A., Walsh, J. & Andrews, P. W. (1998) *Oncogene* **17,** 179–186.
5. Miller, W. H., Jr., Maerz, W. J., Kurie, J., Moy, D., Baselga, J., Lucas, D. A., Grippo, J. F., Masui, H. & Dmitrovsky, E. (1994) *Differentiation (Berlin)* **55,** 145–152.
6. Veal, E., Groisman, R., Eisenstein, M. & Gill, G. (2000) *Oncogene* **19,** 2120–2128.
7. Veal, E., Eisenstein, M., Tseng, Z. H. & Gill, G. (1998) *Mol. Cell. Biol.* **18,** 5032–5041.
8. Di Bacco, A. & Gill, G. (2003) *Oncogene* **22,** 5436–5445.
9. Xu, L., Liu, J. M. & Chen, L. Y. (2004) *J. Hypertens.* **22,** 1579–1587.
10. Kunita, R., Otomo, A. & Ikeda, J. E. (2002) *Genomics* **80,** 456–460.
11. Oka, Y., Rozek, L. M. & Czech, M. P. (1985) *J. Biol. Chem.* **260,** 9435–9442.
12. Sato, Y., Tsuboi, R., Lyons, R., Moses, H. & Rifkin, D. B. (1990) *J. Cell Biol.* **111,** 757–763.
13. Dennis, P. A. & Rifkin, D. B. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 580–584.
14. Ghosh, P., Dahms, N. M. & Kornfeld, S. (2003) *Nat. Rev. Mol. Cell. Biol.* **4,** 202–212.
15. Hendrickson, W. A., Horton, J. R. & LeMaster, D. M. (1990) *EMBO J.* **9,** 1665–1672.
16. Otwinowski, Z. & Minor, W. (1997) *Methods Enzymol.* **276,** 307–326.
17. Terwilliger, T. C. & Berendzen, J. (1999) *Acta Crystallogr. D* **55,** 849–861.
18. Terwilliger, T. C. (2003) *Acta Crystallogr. D* **59,** 45–49.
19. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard. (1991) *Acta Crystallogr. A* **47,** 110–119.
20. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997) *Acta Crystallogr. D* **53,** 240–255.
21. Laskowski, R. A., McArthur, M. W., Moss, D. S. & Thornton, J. M. (1993) *J. Appl. Crystallog.* **26,** 282–291.
22. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997) *Nucleic Acids Res.* **25,** 3389–3402.
23. Holm, L. & Sander, C. (1995) *Trends Biochem. Sci.* **20,** 478–480.
24. Murzin, A. G., Brenner, S. E., Hubbard, T. & Chothia, C. (1995) *J. Mol. Biol.* **247,** 536–540.
25. Rangarajan, E. S., Li, Y., Iannuzzi, P., Tocilj, A., Hung, L. W., Matte, A. & Cygler, M. (2004) *Protein Sci.* **13,** 3006–3016.
26. Reddy, S. T., Chai, W., Childs, R. A., Page, J. D., Feizi, T. & Dahms, N. M. (2004) *J. Biol. Chem.* **279,** 38658–38667.
27. Westlund, B., Dahms, N. M. & Kornfeld, S. (1991) *J. Biol. Chem.* **266,** 23233–23239.
28. Hancock, M. K., Yammani, R. D. & Dahms, N. M. (2002) *J. Biol. Chem.* **277,** 47205–47212.
29. Olson, L. J., Yammani, R. D., Dahms, N. M. & Kim, J. J. (2004) *EMBO J.* **23,** 2019–2028.
30. Byrd, J. C., Park, J. H., Schaffer, B. S., Garmroudi, F. & MacDonald, R. G. (2000) *J. Biol. Chem.* **275,** 18647–18656.
31. York, S. J., Arneson, L. S., Gregory, W. T., Dahms, N. M. & Kornfeld, S. (1999) *J. Biol. Chem.* **274,** 1164–1171.
32. Gordon, P. V., Paxton, J. B. & Fox, N. S. (2005) *J. Endocrinol.* **185,** 265–273.