Solution Structure of a Novel Chromoprotein Derived from Apo-Neocarzinostatin and a Synthetic Chromophore†

Michael D. Urbaniak,‡ Frederick W. Muskett,§ Michael D. Finucane,‡ Stephen Caddick,*,‡ and Derek N. Woolfson‡

Centre for Biomolecular Design and Drug Development, CPES, University of Sussex, Falmer, Lewes Road, Brighton BN1 9QJ, U.K., and Medical Research Council Biomedical NMR Center, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, U.K.

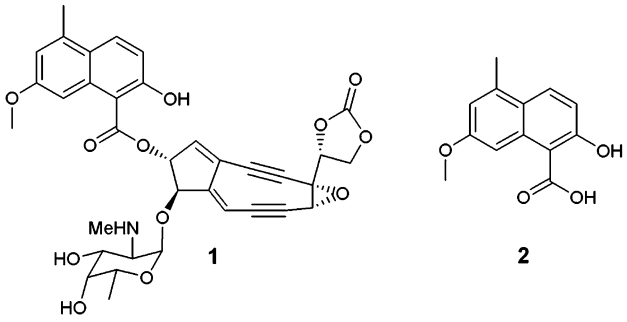
Received May 29, 2002

ABSTRACT: The natural complex Neocarzinostatin comprises a labile chromophore noncovalently bound to an 11.2 kDa protein. We present the first high-resolution structure of a novel complex derived from the recombinant apoprotein bound to a non-natural synthetic chromophore. Fluorescence and nuclear magnetic resonance spectroscopy were used to probe the strength and location of binding. Binding occurred in a location similar to that observed for the chromophore in the natural Neocarzinostatin complex, but with a distinct orientation. These results provide structural evidence that the apoprotein can readily accommodate small druglike entities, other than the natural chromophore within its binding cleft. The clinical use of the natural complex described by others, together with the results reported here, suggests potential applications for small molecule binding by apo-Neocarzinostatin.

The chromoprotein antitumor antibiotic Neocarzinostatin (NCS)[1] has elicited immense interest from chemical, biological, and medical perspectives, and has found clinical application in chemotherapy in Japan against leukemia and cancers of the bladder, stomach, pancreas, liver, and brain (*1*−*3*). NCS comprises a highly reactive chromophore component **1** (Scheme 1) noncovalently bound to an 11 kDa protein (apoNCS) (*4*), and is the most extensively studied member of a class of natural products that includes, among others, the homologous Actinoxanthin (*5*), C-1027 (*6*), Kedarcidin (*7*), and Macromomycin (*8*).

The cytotoxic activity of NCS results from the Bergman-type cycloaromatization of the strained nine-membered enediyne ring of **1** (*9*), producing a diradical responsible for single- and double-stranded DNA cleavage by hydrogen abstraction from the deoxyribose sugar backbone (*10*).

Scheme 1



However, the exact role of the protein remains unclear. ApoNCS alone is not cytotoxic, but provides essential protection for the labile chromophore from heat, UV light, and attack by nucleophiles. Both apo-NCS and holo-NCS enter the yeast *Saccharomyces cerevisiae* by unknown mechanisms (*11*), and penetrate at least as far as the nuclear membrane in mammalian cells (*12*). ApoNCS was once believed to exhibit proteolytic activity against histone H1; however, Heyd et al. showed that highly purified apoNCS does not exhibit such activity (*13*). Conjugation of NCS to poly(styrene-*co*-maleic acid anhydride) produces the pharmaceutical SMANCS, which displays improved tumor uptake and an improved toxicological profile (*14*). The targeting of SMANCS to solid tumors is believed to occur by the enhanced permeability and retention effect, whereby macromolecules leak from poorly formed vasculature surrounding tumors and are retained due to poor lymphatic drainage (*15*).

The purported ability of the protein to act as a transporter for reactive small molecules is significant, and while there are potential difficulties in using protein-based drug delivery systems, its clinical use indicates that such approaches hold promise and are worthy of further investigation. The ability of apoNCS, or variants thereof, to bind to other small

† We are grateful to the EPSRC, BBSRC, AICR, AstraZeneca, GlaxoSmithKline, Novartis, and the Centre for Biomolecular Design and Drug Development at the University of Sussex for supporting our program. We thank the Medical Research Council for NMR spectrometer time at the National Institute for Medical Research.

* To whom correspondence should be addressed: Centre for Biomolecular Design and Drug Development, CPES, University of Sussex, Falmer, Lewes Road, Brighton BN1 9QJ, U.K. Phone and fax: (+44) 1273 678734. E-mail: S.Caddick@Sussex.ac.uk.

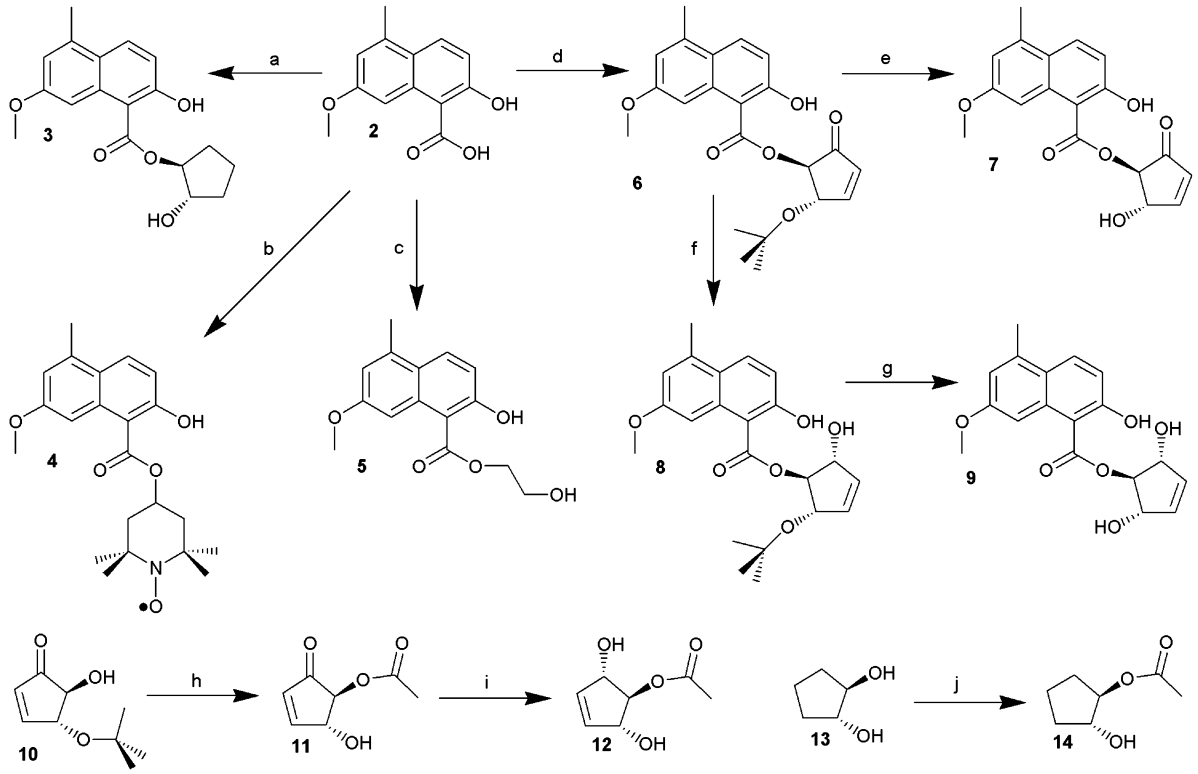
‡ University of Sussex.

§ National Institute for Medical Research.

[1] Abbreviations: apoNCS, apo-Neocarzinostatin; CD, circular dichroism; NCS, Neocarzinostatin; apoNCS^R, recombinant apo-Neocarzinostatin; SMANCS, poly(styrene-*co*-maleic acid anhydride)−NCS conjugate; EtBr, ethidium bromide; hy-TEMPO, 4-hydroxy-2,2,6,6-tetramethylpiperidin-1-oxyl; Cp, *trans*,*trans*-3,4,5-trihydroxycyclopentene; T_m, midpoint unfolding temperature; rmsd, root-mean-square deviation; HSQC, heteronuclear single-quantum coherence; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; TOCSY, total correlated spectroscopy; DQF-COSY, double-quantum-filtered correlation spectroscopy; RP-HPLC, reversed phase high-performance liquid chromatography, MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight.

10.1021/bi0262146 CCC: $22.00
Published on Web 08/17/2002

Scheme 2: Synthesis of Compounds[a]



[a] Reagents and conditions: (a) *trans*-1,2-cyclopentanediol (10 equiv), DCC, CH_2Cl_2, 0 °C, 16 h, 58%; (b) $(COCl)_2$ (1.0 equiv), catalytic DMF, CH_2Cl_2, 0 °C to RT, then Hy-TEMPO (1.0 equiv), NEt_3 (1.0 equiv), RT, 16 h, 11%; (c) ethylene glycol (10 equiv), DCC (1.1 equiv), 1:1 CH_2Cl_2/THF, 0 °C, 16 h, 89%; (d) **10** (1.1 equiv), DCC (1.1 equiv), 1:1 CH_2Cl_2/THF, 0 °C, 16 h, 64%; (e) $TiCl_4$ (2.0 equiv), CH_2Cl_2, −40 °C, 15 min, 67%; (f) $CeCl_3 \cdot 7H_2O$ (1.2 equiv), $NaBH_4$ (1.2 equiv), MeOH, 0 °C, 45 min, 69%; (g) $TiCl_4$ (2.0 equiv), CH_2Cl_2, RT, 15 min, 50%; (h) AcO_2 (2.2 equiv), NEt_3 (2.2 equiv), DMAP (0.3 equiv), CH_2Cl_2, 0 °C, 1 h, 97%; (i) $CeCl_3 \cdot 7H_2O$ (1.4 equiv), $NaBH_4$ (1.4 equiv), CH_2Cl_2, 0 °C, 30 min, 45%; (j) AcO_2 (1.0 equiv), NEt_3 (1.0 equiv), DMAP (0.15 equiv), CH_2Cl_2, 0 °C, 1.5 h, 25%.

molecule drugs may offer opportunities for delivering other entities in vivo. However, such applications would depend on the ability of apoNCS to accommodate molecules other than **1** within the binding cleft.

The three-dimensional structure of NCS shows the naphthoate **2** is deeply buried within the binding cleft of the protein, leading to suggestions that **2** is an important element in the binding of **1** to apoNCS (*16*−*18*). Furthermore, evidence from fluorescence quenching titration experiments shows that analogues of **1** containing **2** bind to apoNCS ($K_d = 10^{-4}-10^{-6}$ M), and altering the substituents of **2** diminishes the extent of binding (*19*). However, currently there is no evidence that these molecules bind within the apoNCS binding cleft, and the additional contacts made by the remainder of the natural chromophore that strengthens binding in the holo complex ($K_d = 10^{-8}-10^{-10}$ M) may further, or completely, define the specific location of binding (*19*, *20*). The possibility that the protein is promiscuous in its binding to small molecule drugs is supported by the reported binding of ethidium bromide (EtBr, $K_d = 10^{-6}$ M) and daunomycin to apoNCS (*21*).

As part of a program examining the potential of apoNCS as a generic drug delivery system, we examined the interaction of small molecules containing the naphthoate **2** with apoNCS in depth. To determine if **2** could be utilized as a motif able to specifically anchor conjugated small molecules within the apoNCS binding cleft, we sought evidence for the location and specificity of binding. This paper presents the first high-resolution structural evidence of the binding of a synthetic chromophore to apoNCS. The structure shows a discrete binding mode and highlights the importance of obtaining detailed structural evidence in this study and similar studies.

MATERIALS AND METHODS

Synthetic Compounds. The naphthoate **2**, the ester **3**, and the cyclopentenone **10** were synthesized as described previously (*19*, *22*−*25*). A series of esters were synthesized from **2** and **10** (Scheme 2).

Production of Recombinant Apo-Neocarzinostatin. A synthetic gene encoding apoNCS was designed taking into account *Escherichia coli* codon usage and constructed from six overlapping oligonucleotide fragments by PCR in a method analogous to that recently reported by Heyd et al. (*13*). The gene was inserted into a modified pCANTAB5 vector (pCANTABB) (*26*), which appended the *lac* promoter, the M13 filamentous phage gIII periplasmic leader sequence, a hexahistidine tag, and a three-amino acid spacer (LQG) 5′ to the apoNCS gene. The recombinant protein was expressed in *E. coli* (HB2151) and purified using affinity chromatography for immobilized Ni(II) ions. Uniform ^{15}N labeling was achieved by growing the bacteria in M9 minimal medium containing $^{15}NH_4Cl$ as the sole nitrogen source.

Characterization of ApoNCS^R. Reversed phase high-performance liquid chromatography (RP-HPLC) and circular dichroism (CD) spectroscopy were performed as described previously (*26*, *27*). All CD spectra were recorded in 25 mM

phosphate (pH 7.0). Proteolysis experiments were conducted by incubating 1 mg/mL calf thymus histone H1 with 0.5 mg/mL apoNCS^R in 50 mM Tris (pH 7.5) at 35 °C, with proteolysis being monitored by SDS−PAGE analysis (*13*).

Fluorescence Spectroscopy. Fluorescence spectra were recorded on a Varian Cary Eclipse fluorescence spectrophotometer fitted with a Peltier temperature control device. Fluorescence quenching titration experiments were conducted using the constant dilution method described by Kondo et al. (*28*) at 25 ± 0.5 °C. Titration experiments used 5 μM naphthoate esters **3**−**5**, **7**, and **9** with 0−150 μM apoNCS^R or 10 μM apoNCS^R with 0−5.0 mM acetate esters **11**, **12**, and **14** in 100 mM sodium acetate buffer (pH 5.0) with 10% (v/v) MeOH. Titration experiments with ethidium bromide (EtBr) used 5 μM EtBr with 0−150 μM apoNCS^R in 100 mM ammonium acetate (pH 5.0). Dissociation constants (K_d) were calculated as the average value of three separate titration experiments analyzed using both a linear (eq 1) (*28*) and a nonlinear method (eq 2) (*13*).

$$\frac{1}{\Delta F} = \frac{K_d}{\alpha[S_T]}\frac{1}{[L_T]} + \frac{1}{\alpha[S_T]} \quad (1)$$

$$\Delta F = \Delta F_{max}\frac{[L_T] - [S_T](\Delta F/\Delta F_{max})}{K_d + [L_T] - [S_T](\Delta F/\Delta F_{max})} \quad (2)$$

where $[S_T]$ and $[L_T]$ are the total concentrations of substrate and ligand, respectively, α is a constant, ΔF is the change in fluorescence, and ΔF_{max} is the change in fluorescence at saturation.

Nuclear Magnetic Resonance Spectroscopy. Nuclear magnetic resonance (NMR) spectroscopy was performed on a Varian Inova 600 MHz or Varian Unity*plus* 500 MHz spectrometer equipped with pulse-field-gradient triple-resonance probes. All data were collected in the phase-sensitive mode using the method of States et al. (*29*) at an operating temperature of 35 °C, with water suppression achieved using the pulsed field gradient based WATERGATE method (*30*). Samples of apoNCS^R were freshly made from lyophilized apoNCS^R in either 100% D_2O or a 10% D_2O/90% H_2O mixture with 25 mM phosphate (pH 5.0), 100 μM TSP, and 0.005% (w/v) sodium azide, to give samples with protein concentrations between 100 μM and 2.2 mM. For samples containing small molecules, additional d_4-MeOH was added to give a final concentration of 10% (v/v) d_4-MeOH.

The three-dimensional $^1H-^{15}N$ experiments were carried out with 2400* (H) × 128* (HN) × 48* (N) complex (*) points, four scans, and a mixing time of 70 ms (TOCSY-HSQC) (*31*), 1800* (H) × 130* (HN) × 50* (N) points, four scans, and a mixing time of 100 ms (NOESY-HSQC) (*32*), 1600* (H) × 84* (HN) × 40* (N) points and eight scans (HNHA) (*33*), and 2000* (H) × 60* (HN) × 28* (N) points and eight scans (HNHB) (*34*). Two-dimensional NOESY spectra were acquired with 4096* × 500* points, 64 scans, and a mixing time of 100 ms. The one-dimensional titration experiments used 250 μM apoNCS in 100 mM sodium acetate (pH 5.0) and 0−2.5 mM EtBr, with no dilution of >5% over the course of the titration. DQF-COSY (*35*) experiments with and without the paramagnetic species were conducted using identical conditions (3200* × 350* points and 48 scans). Two-dimensional $^1H-^{15}N$ HSQC (*36*) experiments required 100 μM uniformly labeled [^{15}N]-apoNCS^R with 200 μM naphthoate ester **3**, **5**, **7**, or **9** or 2.0 mM acetate ester **11**, **12**, or **14** and took 49 min to record (128* × 1200* points and eight scans).

All spectra were processed with the program NMRPipe (*37*) and analyzed on screen with the program XEASY (*38*). Apodization functions and zero filling to the nearest appropriate 2^n data points were applied to the free inductive decays prior to Fourier transformation. A $\pi/2$ shifted sine-bell squared function was used in the direct dimension, and a $\pi/2$ shifted sine-bell squared function with a first point multiplier of 0.5 was used in the indirect dimension(s). The perturbation of backbone amide chemical shifts in the presence of small molecule analogues in 10% (v/v) d_4-MeOH was calculated relative to apoNCS^R recorded in 10% (v/v) d_4-MeOH via $^1H-^{15}N$ HSQC spectra using the minimum chemical shift procedure (eq 3) (*39*, *40*).

$$\min \Delta\delta = \min[(^{HN}\Delta_{ppm})^2 + (^{N}\Delta_{ppm}\alpha_N)^2]^{1/2} \quad (3)$$

where $^{HN}\Delta_{ppm}$ and $^{N}\Delta_{ppm}$ are the $^1H_{HN}$ and $^{15}N_{NH}$ chemical shift changes, respectively, and α_N is a scaling factor to account for the difference in the spectral width of backbone ^{15}N relative to ^1H. An α_N value of $^1/_7$ is used for all residues except glycine ($\alpha_N = {}^1/_5$).

Structure Calculation. Structurally significant intra- and interresidue NOEs were identified in the three-dimensional $^{15}N-^1H$ NOESY-HSQC and two-dimensional ^1H NOESY spectrum of apoNCS^R. The relationship between NOE intensity and interproton distance was calibrated using NOEs corresponding to known distances in regular β-sheet regions of apoNCS. Using this calibration, the NOE intensities were converted to upper distance constraints using the program CALIBA (*41*) with the maximum upper distance limit set to 6 Å. Where appropriate, structural distance corrections were applied to constraints involving methyl and aromatic ring protons (*42*, *43*). In addition, when spectral overlap hindered the determination of volumes for NOE cross-peaks, the distance constraints were set to the upper limit (6 Å). The ratio of diagonal to cross-peak volume in the HNHA spectrum allowed $^3J(H^NH^\alpha)$ coupling constants (±1 Hz) to be determined, which were used to generate backbone ϕ torsional angle constraints.

The high-resolution structures were calculated with the program DYANA, which uses simulated annealing combined with torsional angle dynamics (*44*). To maximize the number of converged structures obtained from an initial set of 100 random starting coordinates, each stage of the calculations included five cycles of the REDAC procedure (*45*). In the final stage of the structure refinement, additional distance constraints were included in the calculation corresponding to NMR-determined hydrogen bonds. The constraints for the two disulfide bonds were progressively included as it became clear that only one possible linkage of a pair of Cys residues was consistent with the structures that were obtained. Hydrogen bonds were included only for residues where an amide proton was detectable after 1 h in D_2O, and where the distance between the hydrogen acceptor and donor atoms was <2.5 Å and the N−H−O bond angle >135°. Hydrogen bonds were included in the calculation using upper and lower distance limits to constrain the separation of the appropriate

NH and O to between 1.8 and 2.3 Å, and the N and O separation to between 2.4 and 3.3 Å.

Structure Analysis. Analysis of the ensemble of solution structures was carried out using the programs MOLMOL (*46*) and DYANA (*44*). In addition, MOLMOL was used to prepare all of the protein structure figures.

RESULTS

Characterization of Recombinant Apo-Neocarzinostatin. Recombinant apo-Neocarzinostatin (apoNCS^R) was produced with an N-terminal hexahistidine tag followed by a three-amino acid spacer (LQG). However, for ease of comparison, the residue numbering used here is identical to that of the wild-type protein. After purification on an immobilized Ni(II) column, RP-HPLC showed a single species and MALDI-TOF MS indicated the correct mature protein product; the high purity of the sample was confirmed by the absence of proteolytic activity against calf thymus histone H1 (*13*). The unusual shape of the CD spectrum of apoNCS^R was in good agreement with that reported for wild-type apoNCS (apoNCS^{WT}) (*47*). ApoNCS^R reversibly unfolded with a midpoint of unfolding (T_m) of 67 °C, as monitored by CD spectroscopy, in good agreement with the T_m of 68 °C reported for apoNCS^{WT} determined using differential scanning calorimetery (*13*).

Structure Determination for ApoNCS^R. Sequence specific ^{1}H and ^{15}N assignments for 105 of 113 residues of apoNCS^R were made using standard procedures (*48*), and have been deposited in the BioMagResBank under accession number 5343 (*49*). The four Pro residues at positions 3, 9, 49, and 105 were not unambiguously identified from the two-dimensional ^{1}H NMR spectra. The remaining unassigned resonances for residues Ala 1, Ala 2, Ala 15, and Arg 82 were not found at previously reported proton shifts (*50*−*52*), nor were they identified elsewhere in the spectra. The nine residues in the apoNCS^R N-terminal hexahistidine tag and spacer, which are likely to be disordered in solution, were not identified. Otherwise, the ^{1}H assignments were in good agreement with values reported by others for apoNCS^{WT} (*50*−*52*).

Solution structures for apoNCS^R were determined from distance and torsional angle restraints listed in Table 1. Coordinates of the final ensemble have been deposited in the Protein Data Bank as entry 1J5H (*53*). An overlay of the 44 structures showed that the residues within β-sheet regions were well defined (Figure 1A), with an average root-mean-square deviation (rmsd) of 0.85 ± 0.19 Å for backbone atoms and 1.59 ± 0.25 Å for heavy atoms. Comparison with the solution structures of apoNCS from other groups was not possible as the atomic coordinates are not publicly available (*54*, *55*).

Comparison of the mean NMR structure with the apoNCS^{WT} crystal structure (*56*) showed good agreement, with an rmsd of 1.87 Å for backbone atoms in secondary structural elements. The NMR structure comprises primarily 11 β-strands: strands A (residues 4−7), B (17−24), C1 (31−36), C2 (37−39), D (45−47), E (53−56), F (62−69), G (72−77), H (82−87), I (94−98), and J (108 and 109) (*57*). A seven-stranded antiparallel β-sandwich is formed by a three-stranded β-sheet (strands A, B, and F) and a four-stranded β-sheet (strands C1, E, I, and J), with a smaller domain formed by two antiparallel β-strands (strands C2 and D and strands G and H) (Figure 1B). The apoNCS binding site is formed by the cleft between strands C1, C2, D, I, and J and enclosed by loops L3 (residues 40−44), L7 (78−81), and L9 (99−107). The positions of loops L3, L7, and L9 were the primary differences between the two structures; consistent with this, these loops had a high local (three-residue) rmsd in the ensemble of NMR structures and high *B*-factors in the crystal structure (*56*) (Figure 2). It has been proposed that these loops have an enhanced flexibility on the basis of NMR-derived relaxation parameters (*58*, *59*). However, the relaxation parameters do not show enhanced flexibility for loop L1 (residues 9−17), which had a high local (three-residue) rmsd in our structure.

Table 1: NMR Constraints and Structural Statistics

parameter	apoNCS^R	apoNCS^R−**9** complex
no. of restraints for structure calculation		
intraresidue NOEs	194	173
sequential NOEs (i, $i+1$)	206	227
medium-range NOEs (i, $i \leq 4$)	129	115
long-range NOEs (i, $i \geq 5$)	303	295
ϕ angle restraints	56	39
hydrogen bonds	52	74
disulfide bonds	12	12
distance constraint violations		
maximum	0.45 ± 0.03 Å	0.42 ± 0.04 Å
total	27.4 ± 1.80 Å	20.5 ± 1.21 Å
Ramachandran plot analysis		
allowed	54%	63%
additionally allowed	33%	30%
generously allowed	10%	4%
disallowed	3%	2%
coordinate precision (rmsd)		
backbone atoms[a]	0.85 ± 0.19 Å	0.73 ± 0.25 Å
heavy atoms[a]	1.59 ± 0.25 Å	1.56 ± 0.27 Å
all backbone atoms	1.54 ± 0.29 Å	1.38 ± 0.27 Å
all heavy atoms	2.02 ± 0.27 Å	2.03 ± 0.30 Å
comparison to crystal structure[b]		
rmsd of backbone atoms[a]	1.87 Å[c]	1.43 Å[d]

[a] Calculated on secondary structural elements: A (residues 4−7), B (17−24), C1 (31−36), C2 (37−39), D (45−47), E (53−56), F (62−69), G (72−77), H (82−87), I (94−98), J (108 and 109). [b] Using the mean NMR structure. [c] Compared to apoNCS (*56*). [d] Compared to NCS (*17*).

Determination of Dissociation Constants. Fluorescence quenching titration experiments were used to determine the dissociation constants (K_d) for the apoNCS^R−small molecule complexes. Initially, two compounds known to bind to apoNCS^{WT} were studied to validate the method. Hirama et al. (*19*) report that both enantiomers of compound **3** bind to apoNCS^{WT} with K_d values of 66 and 77 μM for the (*S*,*S*) and (*R*,*R*) enantiomers, respectively. Our measurement of a K_d of 50 ± 4 μM for the binding of racemic **3** to apoNCS^R was in good agreement. However, we obtained a K_d of 140 ± 22 μM for the binding of ethidium bromide (EtBr) to apoNCS^R, which differed significantly from the literature values of K_d of 1−2 μM (*13*, *21*). Mohanty et al. (*21*) report that binding of EtBr causes perturbation of the chemical shift of the δ-Me protons of Leu 45 visible in one-dimensional ^{1}H NMR spectroscopy, which presents an alternative method for determining the K_d of the EtBr−apoNCS^R complex. Our NMR titration of apoNCS^R with EtBr gave a K_d of 300 ± 70 μM, which was in reasonable agreement with the value we obtained by fluorescence considering the difference in the concentrations that were employed.

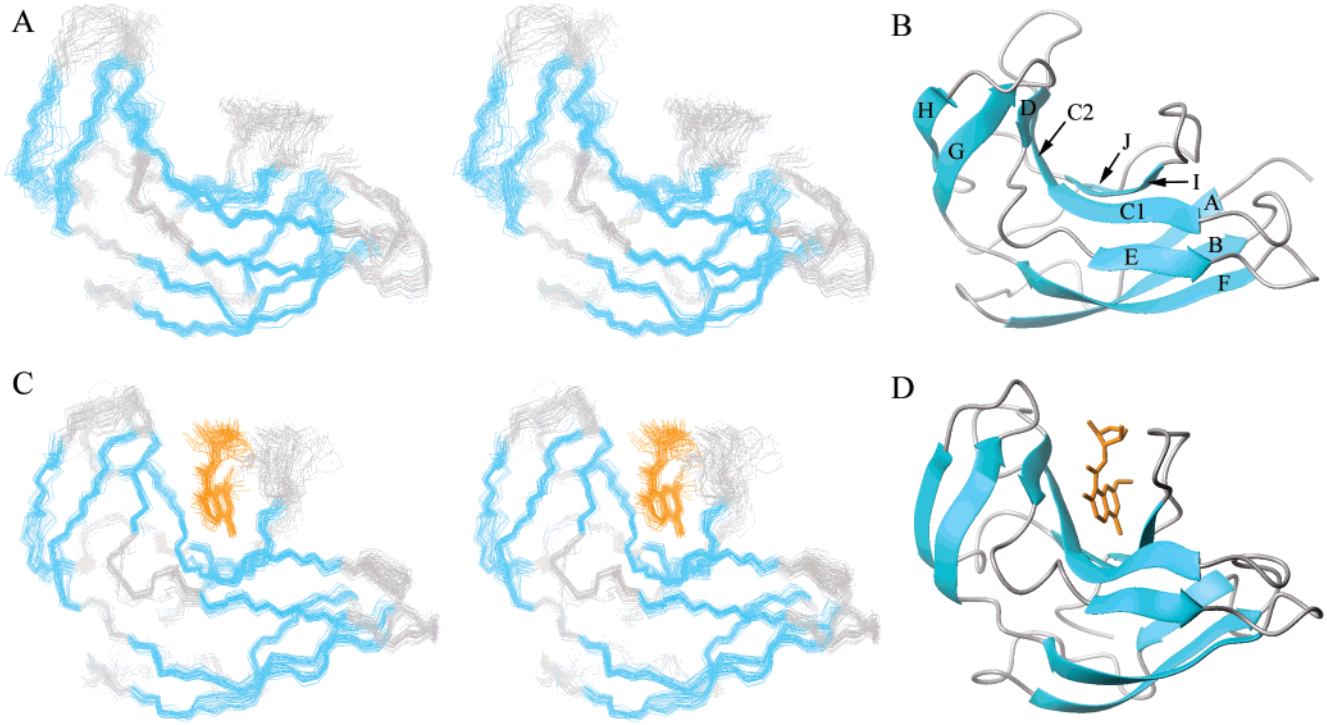


FIGURE 1: Solution structures of apoNCS^R and the **9**−apoNCS^R complex. (A) Stereoview of the 44 lowest-energy structures of apoNCS^R. (B) Ribbon representation of apoNCS with strands labeled (*56*). (C) Stereoview of the 44 lowest-energy structures of the **9**−apoNCS^R complex. (D) Ribbon representation of the **9**−apoNCS^R complex. β-Strands are colored cyan and loops colored gray, and **9** is colored orange.

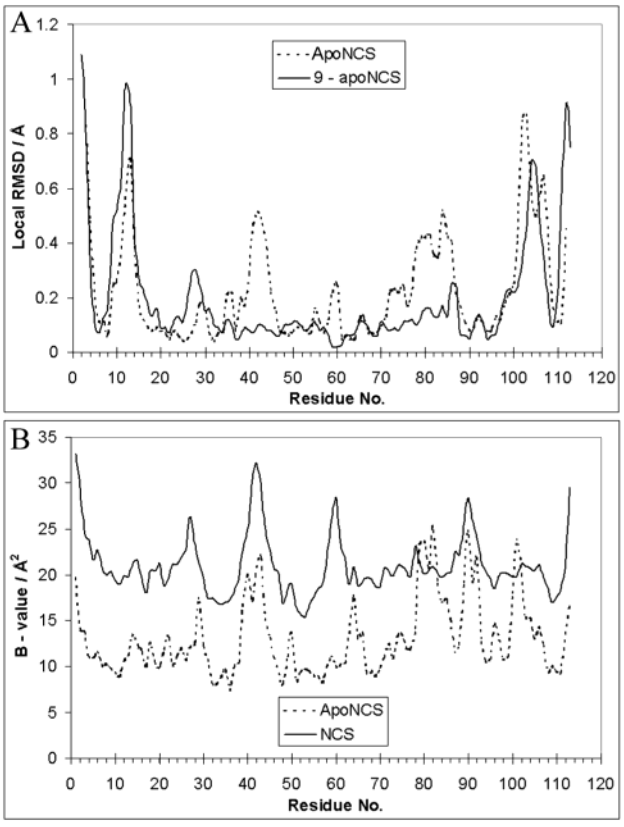


FIGURE 2: Local (three-residue) rmsd for NMR solution structures and crystal structure *B* values. (A) Local (three-residue) rmsd for backbone atoms in the ensemble of solution structures. (B) Crystal structure *B* values for backbone atoms (*17*, *56*).

Changes in the fluorescence of the naphthoate were used to follow the binding of the naphthoate esters, but the low intrinsic fluorescence of the corresponding acetate esters required an alternative approach. Instead, changes to the fluorescence properties of apoNCS^R were used as molecules occupying the binding site of apoNCS perturb Trp 39 on strand C2 (*17*, *21*), whereas the second Trp contributes little to the fluorescence of the protein (*60*). However, any interactions remote from Trp 39 could not be detected. Dissociation constants determined for the compounds under the standard conditions used in this study are reported in Table 2. Of the acetate esters that were tested, only **11** produced a change in the fluorescence of apoNCS^R sufficient for a K_d value to be determined, and displayed significantly weaker binding than the naphthoate esters.

Table 2: Dissociation Constants

compd	K_d (μM)	compd	K_d (μM)	compd	K_d (μM)
3	46 ± 4.1	**5**	260 ± 33	**9**	230 ± 60
4	870 ± 120	**7**	115 ± 10	**11**	11000 ± 3300

Observation of Complex Formation by NMR Spectroscopy. The use of a spin-labeled ligand allows the residues of a protein that are in the proximity of the ligand-binding site to be identified by increased line widths induced in NMR spectra. Phase sensitive DQF-COSY experiments are particularly suited to observing such increases in line width, as antiphase multiplets are severely attenuated even by modest line broadening (*61*). To assess the specificity of binding of spin-labeled ligand **4** to apoNCS^R, DQF-COSY spectra were recorded in the presence and absence of **4**. Data acquisition and spectral processing using identical parameters produced spectra that were clearly distinguishable (Figure 3). No changes to the apoNCS^R spectrum were noted upon addition of an equivalent quantity of the spin-label hy-TEMPO alone

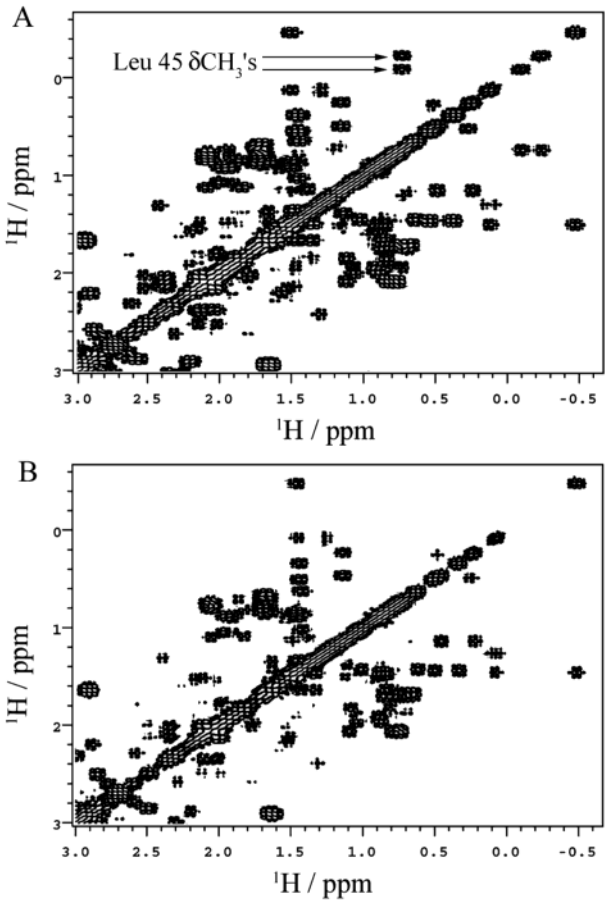


FIGURE 3: Phase sensitive DQF-COSY spectra of (A) apoNCS and (B) apoNCS with 1 equiv of **4**.

to the protein (data not shown), indicating that the naphthoate moiety was responsible for binding to the protein. The δ-CH_3 protons of Leu 45 at -0.23 and -0.13 ppm were clearly attenuated in the presence of **11**, with similar effects observed for Leu 45 Hβ2, Trp 39 Hδ1, and Trp 39 Hξ1 that all form close contact with **1** in the natural holo complex (*17*). Other resonances attenuated upon binding of **11** could not be identified unambiguously as the naphthoate moiety caused chemical shift perturbation that was sufficient to hinder assignment of crowded regions of the spectrum.

Thus, a more general approach to pinpointing small molecule binding to the protein using NMR spectroscopy was sought. Minimum chemical shift perturbation (min $\Delta\delta$) analyses of $^1H-^{15}N$ HSQC spectra have been used in other systems to follow changes in backbone amide resonances upon complex formation by proteins (*39*, *40*). Such chemical shift perturbations can be attributed to the proximity of small molecule binding provided that the protein does not undergo significant structural rearrangement upon complex formation. The location of chemical shift perturbations was visualized by using the min $\Delta\delta$ values of backbone amide resonances to determine the width of the ribbon in a ribbon representation of NCS. Similar perturbations were observed with naphthoate esters **3**, **5**, **7**, and **9** at 2 molar equiv (Figure 4), whereas no significant perturbations were observed with the acetate esters even at 200-fold excess over the protein.

The interactions were clearly localized to the apoNCS binding cleft, with resonances from residues in strands C1, C2, and I being the most strongly perturbed in each complex. These strands form the base and two sides of the binding site, and are in close contact with the naphthoate moiety in the natural holo complex (*17*). Strand D and loops L3, L7, and L9 that surround the apoNCS binding site showed slight to moderate shifts that varied with the nature of the ester group present, and were consistent with the ester moiety occupying the top of the binding site. Despite the clear utility of the method, measurement of the minimum perturbation of chemical shifts provided insufficient evidence to define the orientation in which the naphthoate esters bind.

*Structure Determination for the ApoNCS^R−**9** Complex.* To define the orientation of binding unambiguously, we determined solution structures for an apoNCS^R−naphthoate ester complex; the *meso* compound **9** was selected for this experiment. Sequence specific ^1H and ^{15}N assignments were made for all apoNCS^R residues in the complex, excluding residues 1−3, 9, 15, 49, 82, and 105, and have been deposited in the BioMagResBank under accession number 5344 (*49*). The backbone NH assignments agreed with those obtained using the min $\Delta\delta$ technique, confirming the validity of the use of the min $\Delta\delta$ technique in our study. A large number of significant chemical shift perturbations occurred upon complex formation, including for Phe 78 (HN, Hα, and Hβ), which is believed to be important in chromophore binding and release due to the benzyl side chain adopting different conformations in apo- and holo-NCS (*62*).

The solution structures for the apoNCS^R−**9** complex were determined from the distance and torsional angle restraints listed in Table 1. Coordinates of the final ensemble have been deposited in the Protein Data Bank as entry 1J5I (*53*). No structurally significant intramolecular constraints were identified for **9**, and all of the observed intermolecular NOEs were between apoNCS and the four aromatic ring protons of **9**, leaving the *trans*,*trans*-3,4,5-trihydroxycyclopentene moiety (Cp) unrestrained. An overlay of the 44 structures that were produced showed that the residues within β-sheet regions were well-defined (Figure 1C), with an average rmsd of 0.73 ± 0.25 Å for backbone atoms and 1.56 ± 0.27 Å for heavy atoms. Despite the small number of intermolecular constraints, the orientation of the naphthoate was unambiguously defined at the bottom of the apoNCS^R binding site with an rmsd of 0.54 ± 0.36 Å for all heavy atoms of **9**.

The mean NMR structures for apoNCS^R and the complex were in good agreement with an rmsd of 1.78 Å for backbone atoms in β-sheet regions. The local (three-residue) rmsd of the ensemble was lower in the complex than in apoNCS^R for loops L3 and L7, though loops L1 and L9 showed little improvement (Figure 2A). A similar reduction in the flexibility of the loop regions that surround the binding sight is reported for NCS (*17*, *56*) and C-1027 (*63*) upon chromophore binding (Figure 2B). Loop L9 was closer to strand A in our complex, while loop L7 was drawn closer to strand D, bringing the side chain of Phe 78 further over the top of the cleft.

The mean NMR structure of the **9**−apoNCS^R complex and the crystal structure of NCS (*17*) showed excellent agreement between the structures of the proteins with an rmsd of 1.43 Å for backbone atoms in β-sheet regions. The binding cleft was narrower and shallower in the NMR structure, with the separation between Gly 104 (loop L9) and Ala 38 (strand C2) decreased from 15.5 to 10.8 Å, and loops L3 and L7 displaced to form closer contacts with **9**. The location of **1**

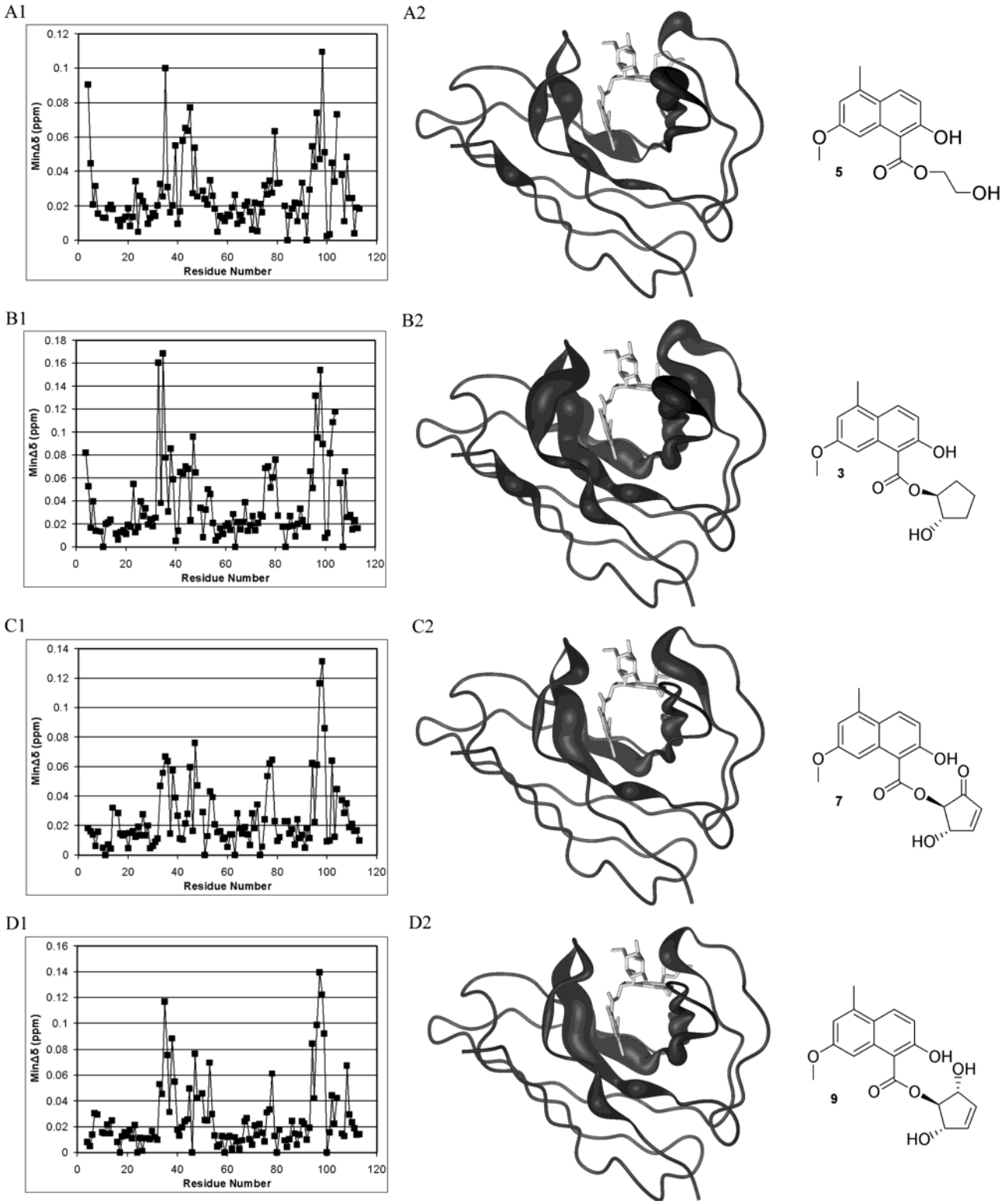


FIGURE 4: Binding of naphthoate esters to apoNCS^R: (A) **5**, (B) **3**, (C) **7**, and (D) **9**. (1) Backbone NH min $\Delta\delta$ analysis vs residue number. (2) Backbone NH min $\Delta\delta$ values mapped onto apoNCS (*17*). The thickness of the ribbon corresponds to four groups: >0.1 ppm, "highly shifted" as 1.6 Å thick; 0.1−0.06 ppm, "moderately shifted" as 1.1 Å thick; 0.06−0.04 ppm, "slightly shifted" as 0.6 Å thick; and <0.04 ppm, "unshifted" 0.2 Å thick. The position of the NCS chromophore (light gray) is shown to aid identification of the binding site.

and **9** within the apoNCS binding pocket, although similar, was distinct in the two structures, with the naphthoate portion of **9** displaced relative to the naphthoate portion of **1** to adopt a more central position in the binding site (Figure 5). The naphthoate of **9** had rotated such that the 7′-OMe group was close to the position occupied by the 2′-OH group of the naphthoate of **1**, destroying the hydrogen bond between the 7′-OMe group and the NH group of Gly 35 seen in the natural complex. The hydrogen bond between the naphthoate carbonyl and Ser 98 γOH (3.1 Å O−O separation) seen in the natural complex was replaced with a hydrogen bond between the 2′-OH group and Ser 98 γOH (3.5 Å O−O separation). The Cys 37−Cys 47 disulfide bond was in closer contact with the naphthoate of **9** than in the natural complex,

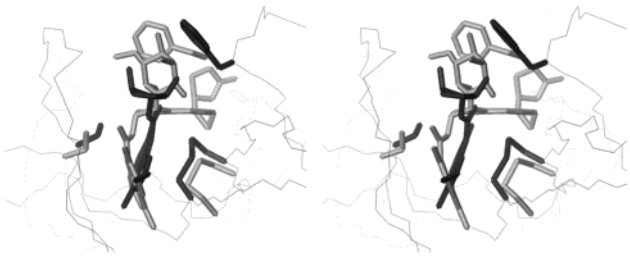


FIGURE 5: Overlay of the **9**−apoNCS^R NMR solution structure with the crystal structure of NCS (*17*). NCS is light gray, and the **9**−apoNCS^R complex is dark gray: top right, Phe 78; bottom left, Ser 98; and bottom right, Cys 37 and Cys 47.

and the Phe 78 side chain was displaced to fit over the Cp portion of **9**. The side chain of Trp 39 and Phe 52 did not lie across either end of the binding pocket as seen in the natural complex, although their positions were consistent with those seen in the apoNCS^R NMR structures. Many common hydrophobic contacts were seen in both the natural complex and our apoNCS^R−**9** complex, including Gly 35, Leu 45, Phe 78, Val 95, Gly 96, and Gly 102.

DISCUSSION

The recombinant apoNCS produced for this study was isolated in a pure, fully folded form free from contaminating proteases. There was excellent agreement between our ensemble of solution structures determined by NMR spectroscopy and the previously reported crystal structure of apoNCS^{WT} (*56*).

A K_d of 220 ± 80 μM for the EtBr−apoNCS^R complex was determined using NMR and fluorescence spectroscopy, which is in poor agreement with literature values determined by fluorescence alone (K_d = 1−2 μM) (*21*). Since the perturbation of a specific binding site residue in the NMR spectra of apoNCS was used to confirm our value, we are confident that the value that was obtained relates specifically to formation of the EtBr−apoNCS^R complex. It may be possible that the tighter binding observed by others is due to contamination of their protein preparations by DNA, which is known to strongly affect the fluorescence of EtBr.

K_d values of 10^{-4}−10^{-5} M were obtained for the naphthoate esters binding to apoNCS^R, while the ester groups alone interacted with the protein only very weakly. Thus, it was the naphthoate **2** that conferred most of the binding to apoNCS^R. However, as binding is only moderate (submillimolar) in comparison to that in the natural ligand (nanomolar), our results confirm, as expected, that the binding stability of the natural complex is not determined by the naphthoate alone.

Our NMR spectroscopic studies indicated that binding occurred within the apoNCS binding site at the same location as occupied by the natural chromophore. Analysis of min $\Delta\delta$ values identified regions of perturbation consistent between the different compounds, corresponding closely to the position of the naphthoate in the holo complex. However, this analysis was insufficient to orient the analogues within the binding site. This could only be obtained by determining a solution structure of the complex. The determination of structures involving weakly bound small molecules is nontrivial. Nonetheless, we report here a structure of the apoNCS^R−**9** complex in which the position of **9** was well-defined, the first high-resolution structure of a synthetic chromophore bound to apoNCS. The differences observed between the position of the naphthoate in our NMR structure and the natural complex showed that not all interactions were conserved: the naphthoate portion of **9** adopted a more central position in the binding site. Thus, the naphthoate **2** should be considered more as a general locator than as an anchor able to confer specific binding to the apoNCS binding cleft.

The naphthoate **2** plays a dual role in the natural complex; not only does **2** bind deeply within the apoNCS binding cleft (*17*), but it also intercalates with DNA (*20*). It can be postulated that due to this dual role **2** may not be optimized for binding to apoNCS, and therefore, synthetic molecules may be found that are able to perform effectively as specific anchors into apoNCS^R. The discrepancy in binding that we observed also highlights the importance of detailed structural studies as a basis for rational drug design; a parsed down ligand should not be expected simply to bind in the same manner as seen in the natural complex.

The promiscuity of interactions within the apoNCS binding cleft holds promise for its application as a generic drug delivery vehicle in that the protein may be able to accommodate a variety of non-natural ligands and thus improve their tumor targeting in vivo. The use of a protein-based delivery system has advantages over the use of synthetic polymer-based formulations as interactions within the protein are well-defined and amenable to structural studies and may be tailored by protein redesign or mutagenesis to modulate the affinity of binding.

REFERENCES

1. Ishida, N., Miyuzaki, K., Kumagai, K., and Rikimaru, M. (1965) *J. Antibiot. 18*, 68−76.
2. Edo, K., Mizugaki, M., Koide, Y., Seto, H., Furihata, K., Otaken, N., and Ishida, N. (1985) *Tetrahedron Lett. 26*, 331−334.
3. Maeda, H., Edo, K., and Ishida, N. (1997) in *Neocarzinostatin: The Past, Present and Future of an Anticancer Drug*, pp 287, Springer-Verlag, Tokyo.
4. Napier, M. A., Holmquist, B., Strydom, D. J., and Goldberg, I. H. (1979) *Biochem. Biophys. Res. Commun. 89*, 635.
5. Khoklov, A. S., Cherches, B. Z., Restiekov, P. D., Smirnova, G. M., Prokoptzena, T. A., Koloditskaya, T. A., and Smirnov, V. V. (1969) *J. Antibiot. 22*, 541−544.
6. Otani, T., Minami, Y., Marunaka, T., Zang, R., and Zie, M.-Y. (1988) *J. Antibiot. 41*, 1580−1585.
7. Lam, K. S., Hesler, G. A., Gustavson, D. R., Crosswell, A. R., Veitch, J. M., Forenza, S., and Tomita, K. (1991) *J. Antibiot. 44*, 472−478.
8. Chimura, H., Ishizuka, M., Hamada, M., Hori, S., Kimura, K., Iwanaga, J., Takeuchi, T., and Umezawa, H. (1968) *J. Antibiot. 21*, 44−49.
9. Myers, A. G. (1987) *Tetrahedron Lett. 28*, 4493−4496.
10. Smith, A. D., and Nicolaou, K. C. (1996) *J. Med. Chem. 39*, 2103−2117.
11. Schaus, S. E., Cavalieri, D., and Myers, A. G. (2001) *Proc. Natl. Acad. Sci. U.S.A. 98*, 11075−11080.
12. Maeda, H., Aikawa, S., and Yamashita, A. (1975) *Cancer Res. 35*, 554−559.
13. Heyd, B., Lerat, G., Adjadj, E., Minard, P., and Desmadril, M. (2000) *J. Bacteriol. 182*, 1812−1818.
14. Maeda, H., Ueda, M., Morinaga, T., and Matsumoto, T. (1985) *J. Med. Chem. 28*, 455−461.
15. Maeda, H. (2001) *Adv. Drug Delivery Rev. 46*, 169−185.
16. Takashima, H., Amiya, S., and Kobayashi, Y. (1991) *J. Biochem. 109*, 807−810.
17. Kim, K. H., Kwon, B. M., Myers, A. G., and Rees, D. C. (1993) *Science 262*, 1024−1046.

18. Tanaka, T., Hirama, M., Fujita, K.-I., Imajo, S., and Ishiguro, M. (1993) *J. Chem. Soc., Chem. Commun.*, 1205−1207.
19. Takahashi, K., Tanaka, T., Suzuki, T., and Hirama, M. (1994) *Tetrahedron 50*, 1327−1340.
20. Povik, L. F., Dattagupta, N., Warf, B. C., and Goldberg, I. H. (1981) *Biochemistry 20*, 4007−4014.
21. Mohanty, S., Seiker, L. C., and Drobny, G. P. (1994) *Biochemistry 33*, 10579−10590.
22. Takahashi, K., Suzuki, T., and Hirama, M. (1992) *Tetrahedron Lett. 33*, 4603−4604.
23. Rucker, M., and Bruckner, R. (1997) *Synlett*, 1187−1189.
24. Caddick, S., Khan, S., Frost, L. M., Smith, N. J., Cheung, S., and Pairaudeau, G. (2000) *Tetrahedron 56*, 8953−8958.
25. Caddick, S., Cheung, S., Doyle, V. E., Frost, L. M., Soscia, M. G., Delisser, V. M., Williams, M. R., Etheridge, Z. C., Khan, S., Hitchcock, P. B., Pairaudeu, G., and Vile, S. (2001) *Tetrahedron 57*, 6295−6303.
26. Finucane, M. D., Tuna, M., Lees, J. H., and Woolfson, D. N. (1999) *Biochemistry 38*, 11604−11612.
27. Finucane, M. D., and Woolfson, D. N. (1999) *Biochemistry 38*, 11613−11623.
28. Kondo, H., Nakatani, M., and Hirami, K. (1976) *J. Biochem. 79*, 393−405.
29. States, D. J., Haberkorn, R. A., and Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286−292.
30. Sklenar, V., Piotto, M., Leppik, R., and Saudek, V. (1993) *J. Magn. Reson. 102*, 241−245.
31. Wijmenga, S. S., Mierlo, C. P. M. v., and Steening, E. (1996) *J. Biomol. NMR 8*, 300−319.
32. Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A., and Bax, A. (1989) *J. Am. Chem. Soc. 111*, 1515−1517.
33. Vuister, G. W., and Bax, A. (1993) *J. Am. Chem. Soc. 115*, 7772−7777.
34. Archer, S. J., Ikura, M., Torchia, D. A., and Bax, A. (1991) *J. Magn. Reson. 95*, 636−641.
35. Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wurtrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479−485.
36. Bodenhausen, G., and Ruben, D. J. (1980) *Chem. Phys. Lett. 69*, 185−189.
37. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277−293.
38. Bartels, C., Xia, T. H., Billeter, M., Guntert, P., and Wurthrich, K. (1995) *J. Biomol. NMR 5*, 1−10.
39. Falmer, B. T., Constantine, K. L., Goldfarb, V., Friedrichs, M. S., Wittekind, M., Yanchanus, J., Robertson, J. G., and Mueller, L. (1996) *Nat. Struct. Biol. 3*, 995−997.
40. Williamson, R. A., Carr, M. D., Frenkiel, T. A., Feeney, J., and Freedman, R. B. (1997) *Biochemistry 36*, 13882−13889.
41. Guntert, P., Braun, W., and Wuthrich, K. (1991) *J. Mol. Biol. 217*, 517−530.
42. Wuthrich, K., Billeter, M., and Braun, W. (1993) *J. Mol. Biol. 169*, 949−961.
43. Wutrich, K. (1986) *NMR of proteins and nucleic acids*, 1st ed., Wiley-Interscience, New York.
44. Guntert, P., Mumenthaler, C., and Wurthrich, K. (1997) *J. Mol. Biol. 273*, 283−298.
45. Guntert, P., and Wutrich, K. (1991) *J. Biomol. NMR 1*, 446−456.
46. Koradi, R., Billeter, M., and Wuthrich, K. (1992) *J. Mol. Graphics 14*, 51−55.
47. Saito, K., Sato, Y., Edo, K., Akiyama-Murai, Y., Koide, Y., Ishida, N., and Mizugaki, M. (1989) *Chem. Pharm. Bull. 37*, 3078−3082.
48. Evans, J. N. S. (1995) *Biomolecular NMR spectroscopy*, Oxford University Press, Oxford, U.K.
49. Seavey, B. R., Farr, E. A., Westler, W. M., and Markley, J. L. (1991) *J. Biomol. NMR 1*, 217−236.
50. Adjadj, E., Mispelter, J., Quiniou, E., Dimicoli, J.-L., Favaudon, V., and L'host, J.-M. (1990) *Eur. J. Biochem. 190*, 263−271.
51. Remerowski, M. L., Glaser, S. J., Sieker, L., Samy, T. S. A., and Drobny, G. P. (1990) *Biochemistry 29*, 8401−8409.
52. Gao, X., and Buckhart, W. (1991) *Biochemistry 30*, 7730−7739.
53. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) *Nucleic Acids Res. 28*, 235−242.
54. Adjadj, E., Quiniou, E., Mispelter, J., Favaudon, V., and Lhoste, J.-M. (1992) *Eur. J. Biochem. 203*, 505−511.
55. Gao, X. (1992) *J. Mol. Biol. 225*, 125−135.
56. Teplyakov, A., Obmolova, G., Wilson, K., and Kuromizu, K. (1993) *Eur. J. Biochem. 213*, 737−741.
57. Kabsch, W., and Sanders, C. (1983) *Biopolymers 22*, 2577−2637.
58. Mispelter, J., Lefevre, C., Adjadj, E., Quiniou, E., and Favaudon, V. (1995) *J. Biomol. NMR 5*, 233−244.
59. Izadi-pruneyre, N., Quiniou, E., Blouquit, Y., Perez, J., Minard, P., Desmedril, M., Mispelter, J., and Adjadj, E. (2001) *Protein Sci. 10*, 2228−2240.
60. Edo, K., Saito, K., Matsuda, Y., Akiyama-Murai, Y., Mizugaki, M., Koide, Y., and Ishida, N. (1991) *Chem. Pharm. Bull. 39*, 170−176.
61. Petros, A. M., Mueller, L., and Kopple, K. D. (1990) *Biochemistry 29*, 10041−10048.
62. Imajo, S., Ishiguro, M., Tanaka, T., Hirama, M., and Teplyakov, A. (1995) *Bioorg. Med. Chem. 3*, 429−436.
63. Tanaka, T., Fukuda-Ishisaka, S., Hirama, M., and Otani, T. (2001) *J. Mol. Biol. 309*, 267−283.

BI0262146