Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 11578–11583, September 1998
Biochemistry

Reactivity of the HIV-1 nucleocapsid protein p7 zinc finger domains from the perspective of density-functional theory

A. T. MAYNARD*, M. HUANG, W. G. RICE, AND D. G. COVELL*

Frederick Cancer Research and Development Center, National Cancer Institute, Science Applications International Corporation, Frederick MD 21702

Communicated by Robert G. Parr, University of North Carolina, Chapel Hill, NC, July 15, 1998 (received for review March 17, 1998)

ABSTRACT **The reaction of the human immunodeficiency virus type 1 (HIV-1) nucleocapsid protein p7 (NCp7) with a variety of electrophilic agents was investigated by experimental measurements of Trp^{37} fluorescence decay and compared with theoretical measures of reactivity based on density-functional theory in the context of the hard and soft acids and bases principle. Statistically significant correlations were found between rates of reaction and the ability of these agents to function as soft electrophiles. Notably, the molecular property that correlated strongest was the ratio of electronegativity to hardness, χ^2/η, a quantity related to the capacity of an electrophile to promote a soft (covalent) reaction. Electronic and steric determinants of the reaction were also probed by Fukui function and frontier-orbital overlap analysis in combination with protein–ligand docking methods. This analysis identified selective ligand docking regions within the conserved zinc finger domains that promoted reaction. The Cys^{49} thiolate was found overall to be the NCp7 site most susceptible to electrophilic attack.**

Human immunodeficiency virus type 1 (HIV-1) encodes Gag and Gag-Pol precursor polyproteins, which are processed to form several mature proteins [(matrix p17, capsid p24, nucleocapsid p7 (NCp7), protease (PR), reverse transcriptase (RT), and integrase (IN)] (1). The emergence of PR and RT drug-resistant HIV strains (2) underscores the importance of pursuing new antiviral strategies. The two retroviral zinc fingers of NCp7 offer an alternative target that may not demonstrate cross-resistance with existing drugs. NCp7 is required for the selection and packaging of the viral genome, as well as additional functions critical to viral replication (3). The core of each finger is composed of one His and three Cys residues, tetrahedrally coordinated to Zn^{2+}, to form a peptide unit C-X_2-C-X_4-H-X_4-C (CCHC). This motif is absolutely conserved among all known strains of retroviruses, except human foamy viruses (4–6). Thus, the NCp7 zinc finger domains represent targets that are potentially nonpermissive to mutations. Various electrophilic agents have been identified that cause HIV-1 inhibition by covalent modification of the nucleophilic zinc finger Cys thiolates (7–12). Notably, the two NCp7 zinc fingers, as well as each cysteine, are not equally vulnerable to electrophilic attack. In the case of disulfide agents 2,2′-dithiobisbenzamides (DIBAs) and 2,2′-dithiopyridine, reaction occurs preferentially at the C-terminal zinc finger (13–15). Moreover, Cys^{49} was found to be substantially more reactive than the other CCHC cysteines (15). Recent results also demonstrated that some agents selectively react with the NCp7 zinc fingers without affecting other cellular zinc finger proteins (16).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

0027-8424/98/9511578-6$0.00/0
PNAS is available online at http://www.pnas.org.

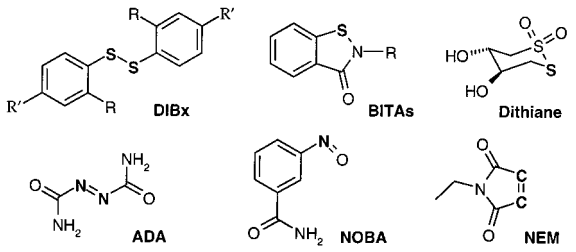

FIG. 1. Electrophilic chemotypes that react with the NCp7. Electrophilic sites are **boldface**. DIBx compounds: DIBA-1 (R = -CONH-phenyl-SO_2NH_2, R′ = H), DIBA-2 (R = -CONH-phenyl-$SO_2NHCOCH_3$, R′ = H), DIBA-3 (R = -$CONH_2$, R′ = -$NHCOCH_3$), DIBPH (R = OH, R′ = H), DIBTL (R = CH_3, R′ = H). BITA compounds: BITA-1 (R = -phenyl-SO_2NH_2), BITA-1E (R = -$(CH_2)_2$-phenyl-SO_2NH_2), BITA-2 (R = -phenyl-$SO_2NHCOCH_3$).

In this study we focus on criteria that determine electrophilic attack of the NCp7 zinc fingers. The reaction is assumed to be initiated by a complementary electronic response between zinc finger and electrophile, requiring sufficient redox properties and geometrical proximity of the reactants, as related to the interaction of frontier orbitals (FOs). If sterically allowed, the relative electronic interaction of the reactants strongly influences the energetics of the transition state, as developed in the hard and soft acids and bases principle (HSAB) (17). The physical underpinning of HSAB is founded in density-functional theory (DFT) (18–20). Thiolates are substantially softer (polarizable) nucleophiles than other biological metal-chelators (21, 22), with lower-lying vacant FOs. Thus, thiolate reactivity is associated with covalent reaction mechanisms and shows a bias towards soft electrophiles (23); "soft reacts with soft." If a soft–soft mechanism underlies the reactivity of the NCp7 zinc fingers, this may be revealed by their reactivity toward a diverse set of electrophiles. Herein, 12 compounds, constituting six different chemotypes (Fig. 1) were selected for study: DIBAs-1, -2, and -3, 2,2′-dithiobisphenol (DIBPH), 2,2′-dithiobistoluene (DIBTL), benzoisothiazolones (BITAs-1, -2, and -1E), azodicarbonamide (ADA), *cis*-1,2-dithiane-4,5-diol-1,1-dioxide (dithiane), *N*-ethylmaleimide (NEM), and 3-nitrosobenzamide (NOBA). Excluding NEM, members of each chemotype have shown anti-HIV activity by means of NCp7 inhibition (7–10). ADA is being evaluated in Europe for patients with advanced AIDS (24). The reaction

Abbreviations: NCp7, nucleocapsid protein p7; DIBA, 2,2′-dithiobisbenzamide; DIBPH, 2,2′-dithiobisphenol; DIBTL, 2,2′-dithiobistoluene; BITA, benzisothiazolone; ADA, azodicarbonamide; dithiane, *cis*-1,2-dithiane-4,5-diol-1,1-dioxide; NEM, *N*-ethylmaleimide; NOBA, 3-nitrosobenzamide; FO, frontier orbital; HSAB, hard and soft acids and bases principle; DFT, density-functional theory; HOMO-LUMO, highest occupied–lowest unoccupied molecular orbital; EA, electron affinity; IP, ionization potential; AO, atomic orbital; DNP, double numerical polarization; a.u., atomic units.

*To whom reprint requests should be addressed at: LECB-Molecular Structure Section, Science Applications International Corporation, Biomedical Supercomputing Laboratory, Bldg. 430, Room 215, SAIC, FCRDC, NCI, Frederick, MD 21702. e-mail: maynarda@fconyx.ncifcrf.gov or covell@faxpdc.ncifcrf.gov.

rate of NCp7 is derived from the fluorescence decay of Trp^{37} (W^{37}), located in the C-terminal finger. The present study uses a method that corrects for external fluorescence quenching. Electrophile properties are investigated in relation to a wide range of observed reaction rates, and the local nucleophilic properties of the NCp7 zinc fingers are computationally probed to determine the most reactive NCp7 sites. Protein–ligand docking and FO overlap analysis are combined to determine the accessibility of these sites to docking arrangements that promote reaction.

DFT Measures of Reactivity. Two quantities central to DFT, the electronic chemical potential (μ) and hardness (η) (18, 19), measure the electronic energy (E[v, N]) response of a system to electron transfer and redistribution, initiated by a chemical reaction. The linear response is μ,

$$\mu = -\chi = \left(\frac{\partial E}{\partial N}\right)_v \simeq -\frac{\mathrm{IP} + \mathrm{EA}}{2}, \qquad [1]$$

or the negative of electronegativity (χ), where N is the number of electrons, and $v(r)$ is the external potential. In practice μ can be calculated in terms of the ionization potential (IP) and electron affinity (EA) by finite difference. Physically, μ corresponds to the capacity of a system to donate electron density and the electron transfer between reactants flows from high to low μ. The curvature corresponds to the system's hardness,

$$\eta = \frac{1}{2}\left(\frac{\partial^2 E}{\partial N^2}\right)_v \simeq \frac{\mathrm{IP} - \mathrm{EA}}{2}, \qquad [2]$$

which measures the resistance to charge redistribution (18, 19). η corresponds to the highest occupied–lowest unoccupied molecular orbital (HOMO-LUMO) energy gap. Chemical softness (σ) is defined as the inverse of η, $\sigma = 1/\eta$. When an electron acceptor (A) and donor (B) react, there is a net redistribution of electrons from B to A ($\mu_B > \mu_A$). As the product AB is formed, μ_A and μ_B equilibrate such that $\mu_A = \mu_B = \mu_{AB}$ (25). Using this condition, Parr and Pearson (18) derived a second-order energy term,

$$\Delta E_{\mathrm{covalent}} = -\frac{1}{4}\frac{(\mu_B^0 - \mu_A^0)^2}{\eta_A + \eta_B}, \qquad [3]$$

that stabilizes the reaction and underlies the soft–soft interaction of HSAB. This term is attributed to the onset of covalent bonding and becomes amplified with increasing reactant σ and difference in μ. Of interest here is whether this result may influence the reactivity of a series of electrophiles {A}, with a given B, by regulating the activation energy. Furthermore, in the regime $\chi_A^0 \gg \chi_B^0$, the dependence of Eq. **3** on {A} scales as

$$\Delta E_{\mathrm{covalent}}^{A} \simeq -\frac{(\chi_A^0)^2}{\eta_A}, \qquad [4]$$

treating terms involving B as constants.† Therefore $(\chi_A^0)^2/\eta_A$ may be an important determinant of reactivity, and we interpret this quantity as the capacity of electrophile A to promote (stabilize) a covalent (soft) reaction.

The spatial determinants of reaction require a local analysis. The differential change in μ contains a global (η) and local component (f): $d\mu = 2\eta dN + \int f(\vec{r})dv(\vec{r})d\vec{r}$, where $f(r) = (\partial\rho/\partial N)_v$ is the Fukui function (19). Parr and Yang (26) proposed that, early in a reaction, the preferred reaction site corresponds to the most concentrated region of $f(r)$. $f(r)$ measures the electron density $\rho(r)$ response to electron addition (electrophile) $f^+(r)$, or electron donation (nucleophile) $f^-(r)$, when probing a reaction mechanism (27, 28). By electron population analysis, $f(r)$ can be condensed to atoms, providing atom-centered reactivity indices f_k: $f_k^+ = q_k(N+1) - q_k(N)$, $f_k^- = q_k(N) - q_k(N-1)$, where q_k is the charge of atom k. Scaling f_k by σ, $\sigma_k = \sigma f_k$, defines local softness indices.

†Reference A to B: $\chi_A^0 = \alpha\chi_B^0$, $\eta_A = \beta\eta_B$. Substitution in Eq. **3** and multiplication by $4\eta_B^0/(\chi_B^0)^2$ yields $-\frac{\alpha^2}{\beta}\frac{(1-1/\alpha)^2}{1+1/\beta} = -\frac{\alpha^2}{\beta}C(\alpha,\beta)$. As $\beta \gg 1$ and particularly $\alpha \gg 1$, $C(\alpha,\beta) \to 1$. Physical considerations, supported by calculations, suggest an intermediate regime for this study: $\alpha > 1$, $\beta > 1$, thus, $0 < C(\alpha,\beta) < 1$.

METHODS

Experimental Reaction Rates of HIV-1 NCp7. Observed reaction rates of recombinant HIV-1 NCp7 (55 residues) were obtained from fluorescence measurements of W^{37}. Upon reaction, W^{37} fluorescence is diminished by increased internal quenching due to a NCp7 conformational change (8, 14). Time course data were obtained with a Shimadzu RF5000 spectrofluorimeter at 25°C, $\lambda_{EX} = 281$ nm and $\lambda_{EM} = 355$ nm. NCp7 was prepared in 10 mM sodium phosphate buffer (pH 7.0) at 0.3 to 1.5 μM with the electrophile concentration in at least 10-fold molar excess. Chemicals were obtained from the National Cancer Institute repository; Octamer Research Foundation, Mill Valley, CA (NOBA); Parke–Davis Pharmaceuticals, Ann Arbor, MI (DIBA-1, -2, and -3, BITA-1, and -2), and Y. Song Laboratory of Chemical Biology, National Cancer Institute and J. Inman National Institute of Allergy and Infectious Diseases, Bethesda, MD (DIBA-2, BITA-2, and BITA-1E). Stock solutions were prepared in dimethyl sulfoxide.

W^{37} fluorescence was often attenuated by external (nonreactive) quenching due to electrophile Q. Stern–Volmer measurements were used to determine the [Q]-dependence of external quenching of *free* tryptophan. This established appropriate concentrations for an adequate W^{37} signal and provided a means to correct for external quenching in computing the reaction rates. For external quenching only (no reaction), the ratio of the observed fluorescence intensity, I_{obs}, to the [Q] = 0 intensity, I_0, is the Stern–Volmer equation: $I_{obs}/I_0 = 1/(1 + K_Q[Q])$, where $K_Q = k_Q/(k_f + k_s)$, k_Q is the external quenching rate of free tryptophan by Q, k_s is the sum of all other quenching processes, and k_f is the emission rate. For the case of *reacted* NCp7, the ratio is $I_{obs}/I_0 = I_\infty/I_0 = 1/(1 + K_r + K'_Q[Q]) \simeq 1/(1 + K_r)$, where $K_r = k_r/(k_f + k_s)$, k_r is the rate of internal quenching of W^{37} due to reaction, and K'_Q is the Stern–Volmer coefficient of reacted NCp7. The condition $K_r > K_Q[Q] \gg K'_Q[Q]$ was observed—i.e., the decay end-point (I_∞) was independent of external quenching, as verified by adding excess Q following completion of reaction. Given these boundary conditions and assuming pseudo-first-order kinetics, the corrected intensity is $I(t) = A(t)I_{obs}(t)$, where $A(t) = 1 + K_Q[Q]e^{-k_{eff}t}$ is the attenuation correction for external quenching, and k_{eff} is the effective reaction rate ($k_{eff} = k_{rxn}[Q] \simeq k_{rxn}[Q]_0$). Thus, the form of W^{37} decay used in computing the reaction rate constants k_{rxn} was

$$I_{obs}(t) = \frac{[(1 + K_Q[Q])I_{obs}(0) - I_\infty]e^{-k_{eff}t} + I_\infty}{1 + K_Q[Q]e^{-k_{eff}t}}, \qquad [5]$$

which reduces to simple exponential decay as $K_Q \to 0$. Nonlinear least-squares regression was used to fit Eq. **5** to the kinetics data, with k_{eff} and I_∞ as the two fitting parameters.

Evaluating Ligand and Zinc Finger Reactivity. Dmol3.0 (Biosym/MSI, San Diego) was used in evaluating molecular reactivity properties. The double numerical polarization (DNP) basis (29) was augmented with diffuse valence orbitals (DNP+) obtained from atomic calculations (Datom utility) with partial negative ionic charges (C 0.5e, N 0.6e, O 0.7e, S 0.8e, and Zn 0.5e). Thus, the DNP+ basis was double-

zeta quality, including diffuse orbitals on heavy atoms and polarization functions on all atoms. Comparisons of EA, IP, E[v, N], and $\rho(r)$, were made between the DNP+ and DNP basis to ensure basis set stability. The local Vosk–Wilk–Nusair (VWN) exchange-correlation functional (30) was used in the self-consistent field (SCF) solution of the Kohn–Sham equations with frozen inner-core atomic orbitals (AOs). A gradient correction to E[v, N] was calculated using the Becke–Perdew (BP) functional (31, 32). $f(r)$ was condensed to atoms by Hirshfeld partitioning. Ligand EA was computed by EA $= (E_N - E_{N+1})_v$, while the IP was approximated by IP $\simeq \varepsilon^{KS}_{HOMO}$. μ and η were calculated by Eqs. **1** and **2**. Ligand conformations were obtained by sampling the lowest energy geometries derived from molecular dynamics and minimization using the CVFF force field of Discover95.0 (Biosym/MSI). The lowest energy conformer was HF/STO-3G optimized (heavy-atom internal coordinates) using Gaussian-94 (Rev. C.3, Gaussian, Pittsburgh). DFT calculations of the NCp7 zinc finger domains were based on the NMR structure (33). Each 56-atom $[(CCHC)Zn]^{-1}$ coordination complex defined the individual zinc finger units. Larger portions of the zinc finger domains were additionally modeled. The N-terminal finger (finger 1) was treated as Lys^{14}–Pro^{31} (186 atoms) and similarly the C-terminal finger (finger 2) was treated as Gly^{35}–Thr^{50} (174 atoms); residues with side chains distant from the $[(CCHC)Zn]^{-1}$ locus were converted to G. The molecular docking methods have been previously described (16, 34). Each ligand was docked against finger 1 (Arg^{10}–Arg^{34}) and finger 2 (Arg^{32}–Asn^{55}), yielding a family of docking arrangements, ranked in terms of their predicted free energy of binding (noncovalent), ΔG_{bind}. To qualitatively assess their geometrical predisposition to reaction, the FO overlap, $\langle \psi^{ligand}_{LUMO} | \psi^{NCp7}_{HOMO} \rangle$, was evaluated (16). This is an alternative to treating the changing electronic and structural features of ligand and target during the docking process, a calculation not yet feasible.

RESULTS AND DISCUSSION

Correlation of Observed Reaction Rates. Normalized time courses of the NCp7 W^{37} fluorescence decay are shown in Fig. 2. The observed rate constants (k_{rxn}) span three orders of magnitude (Table 1). The model decay of Eq. **5** fit the observed data quantitatively: the average residual standard deviation (RSD) from the normalized decays was 0.005 ± 0.003 relative fluorescence units (RFU) over all time courses. The time course of ADA deviated the most from the model prediction (RSD $= 0.014 \pm 0.003$ RFU) due to systematic deviations

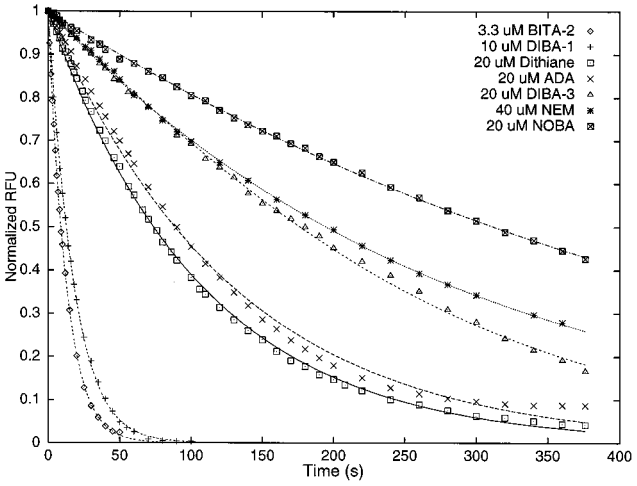

FIG. 2. Experimental NCp7 W^{37} fluorescence decays for the classes of electrophiles studied, expressed in normalized relative fluorescence units (RFU), $I_{norm}(t) = [I_{obs}(t) - I_\infty]/[I(0) - I_\infty]$, compared with model decays, Eq. **5**.

in the early and late portions of the decay, also observed to a lesser degree in the dithiane kinetics. The origin of these deviations was attributed to NCp7 adsorption to the quartz cell (35), though competing reaction pathways could not be ruled out. DFT calculations also showed substantial variation in the ligand electronic properties, where local (σ_n) and global (σ) softness, electronegativity (χ), electron affinity (EA), and combination (χ^2/η) from Eq. **4**, are summarized in Table 1. Together, these parameters define a generic reactivity profile. From Table 1, at least three groups of ligands shared similar profiles: DIBA-1 and DIBA-2; ADA, dithiane, and NOBA; and BITA-1 and BITA-2; this similarity was reflected in concomitant clustering of their observed reaction rates.

Of the properties calculated, χ^2/η correlated most strongly with the observed reaction rates (Fig. 3). A significant correlation existed not only within the DIBx family, the largest chemotype set represented, but also across all chemotypes. Among the DIBx ligands, the linear correlation coefficient between $\ln(k_{rxn})$ and χ^2/η was $r^2 = 0.99$, with probability $P = 4 \times 10^{-4}$. Given the form of Eqs. **3** and **4**, it was of interest to examine correlations with χ and $\sigma = 1/\eta$, separately. These correlations were $r^2 = 0.96$ ($P = 0.003$) for χ and $r^2 = 0.67$ ($P = 0.1$) for σ. Notably, σ alone correlated only qualitatively with the reaction rates. High σ does not

Table 1. Ligand reactivity parameters, observed NCp7 reaction rates, and Stern–Volmer coefficients

Ligand	σ_n^{+}*	$\sigma^{\dagger}$	$\chi^{\ddagger}$	EA§	χ^2/η¶	k_{rxn}‖	K_Q**
Dithiane	1.9 (S—SO_2)	10.59	0.1461	0.0516	0.2260	4.2×10^{-4}	0
ADA	1.8 (N═N)	10.82	0.1451	0.0527	0.2278	4.1×10^{-4}	0.003
NOBA	2.3 (N═O)	13.24	0.1296	0.0541	0.2223	1.5×10^{-4}	0.012
NEM	1.1 (C═C)	9.32	0.1354	0.0282	0.1710	8.8×10^{-5}	0.001
DIBA-1	0.6 (S—S)	13.34	0.1483	0.0734	0.2935	6.7×10^{-3}	0.019
DIBA-2	0.4 (S—S)	12.69	0.1563	0.0774	0.3098	2.6×10^{-2}	0.071
DIBA-3	1.1 (S—S)	12.99	0.1218	0.0448	0.1927	2.9×10^{-4}	0.042
DIBPH	1.5 (S—S)	10.28	0.1249	0.0276	0.1603	7.1×10^{-5}	0.007
DIBTL	1.4 (S—S)	10.16	0.1057	0.0073	0.1136	6.4×10^{-6}	0
BITA-1	0.6 (S)	11.11	0.1328	0.0429	0.1960	1.5×10^{-2}	0.023
BITA-2	0.4 (S)	12.36	0.1348	0.0540	0.2247	2.6×10^{-2}	0.021
BITA-1E	0.7 (S)	11.40	0.1248	0.0371	0.1776	2.1×10^{-4}	0.001

*Local softness ($\sigma_n = \sigma f_n$) of probable ligand reactive site [atomic units (a.u.)].
†Ligand global softness (a.u.), Eq. **2**.
‡Electronegativity (a.u.), Eq. **1**.
§Electron affinity (a.u.).
¶Covalent (soft) reaction capacity, Eq. **4** (a.u.).
‖Observed reaction rate constants ($s^{-1} \cdot \mu M^{-1}$), Eq. **5**, $k_{rxn} = k_{eff}/[Q]_0$.
Stern–Volmer coefficients (μM^{-1}) used in Eq. **5.

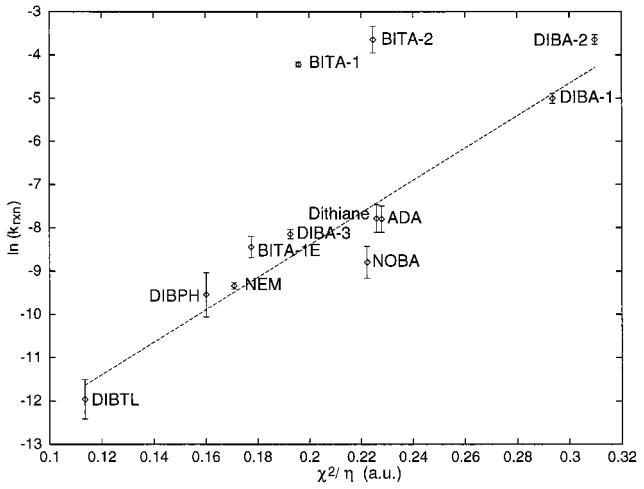

FIG. 3. Correlation between observed NCp7 reaction rates and the ligand capacity to promote a covalent (soft) reaction, χ^2/η, Eq. **4**.

necessarily imply a large χ to drive the reaction, where there is a predicted quadratic dependence on χ in Eq. **3**. However, when combined with χ, σ enhanced the overall correlation of χ^2/η with the experimental data. Thus, the form of Eqs. **3** and **4** was supported. Similarly, simply using EA, the forward linear response of Eq. **1**, $(\partial E/\partial N)_v^+ \simeq \mathrm{EA}$, yielded strong correlations. Correlations of $\ln(k_{\mathrm{rxn}})$ with EA^2/η and EA were $r^2 = 0.95$ ($P = 0.005$) and $r^2 = 0.99$ ($P = 4 \times 10^{-4}$), respectively. Interestingly, local softness was anticorrelated with reaction rate. This suggests that redistribution of $\rho(r)$ throughout the ligand, rather than a local site, stabilized the reaction. Reactivity trends within the DIBx family can largely be understood by the ability of phenyl ring substituents to inductively and/or resonantly withdraw electron density from the disulfide bond. For instance, the inductive effect of *o*-substitution of -OH for $-CH_3$ in comparing DIBTL to DIBPH was present. Moreover, acetylation of the DIBA-1 amide tail to form DIBA-2 had a measurable effect on reactivity, even though the acetyl group is far removed from the disulfide bond (through-bond distance $\approx$ 15 Å). This finding suggests resonance coupling through the full length of the DIBA tails.

The smallest ligands (ADA, dithiane, NOBA, and NEM) followed the same relative and absolute trend in reactivity as the DIBx family (Fig. 3), though each ligand constitutes a different chemotype. ADA and dithiane had nearly identical values of χ^2/η and k_{rxn}. The diminished reactivity of NEM, the hardest electrophile, correlated with decreased χ^2/η. However, the reactivity of NOBA, the softest of the small ligands, was lower than predicted. Although qualitatively correlated with χ^2/η, the BITA compounds deviated significantly from the pattern observed for the other chemotypes. Both BITA-1 and BITA-2 were predicted to have diminished χ and σ relative to their dimeric forms (DIBA-1, DIBA-2), yet their rate constants were comparable. Docking analysis did not reveal significantly enhanced binding energies to account for this discrepancy, although the BITAs achieved closer reactive proximities to the zinc finger thiolates than did the DIBAs. Notably, BITA-1 and -2 have large dipole moments, 8.9 and 12.8 debye, respectively, twice that of the other chemotypes, oriented in the direction of the electropositive thiazolone sulfur. DFT calculations show electronegative regions at the zinc finger Cys thiolates with large local dipoles oriented along the Zn–His axes. Therefore there may be a significant electrostatic component to the BITA reactivity, not treated in our analysis (19).

Excluding BITA-1 and BITA-2, correlations over all compounds with χ^2/η, EA^2/η, χ, and EA were $r^2 = 0.93$ ($P = 7 \times 10^{-6}$), 0.87 ($P = 8 \times 10^{-5}$), 0.76 ($P = 1 \times 10^{-3}$), and 0.90 ($P = 3 \times 10^{-5}$), respectively. The combination χ^2/η gave the strongest correlation. The stronger correlation of EA versus χ is likely due to the greater sensitivity of EA, compared to IP, in measuring electrophilic response. Notably, the calculated IPs of all compounds were similar, in contrast to the large variance in EA. The correlation with χ^2/η suggests that values of χ and η for NCp7 can be empirically estimated by means of Eq. **3**, assuming that the activation energy ($E_{\mathrm{A}}^{\ddagger}$) for reaction of NCp7 with electrophile A has the form $E_{\mathrm{A}}^{\ddagger} = E_{\mathrm{A}}^{*} + \zeta \Delta E_{\mathrm{covalent}}$, where E_{A}^{*} is the residual activation energy and ζ scales the contribution from Eq. **3**.§ The least-squares values (a.u.) of NCp7 were $\chi_{\mathrm{p7}} = 0.031 \pm 0.029$, $\eta_{\mathrm{p7}} = -0.014 \pm 0.020$, with $\zeta = 0.16 \pm 0.03$, yielding a correlation of $r^2 = 0.93$ ($P = 7 \times 10^{-6}$). Similarly, imposing the con-

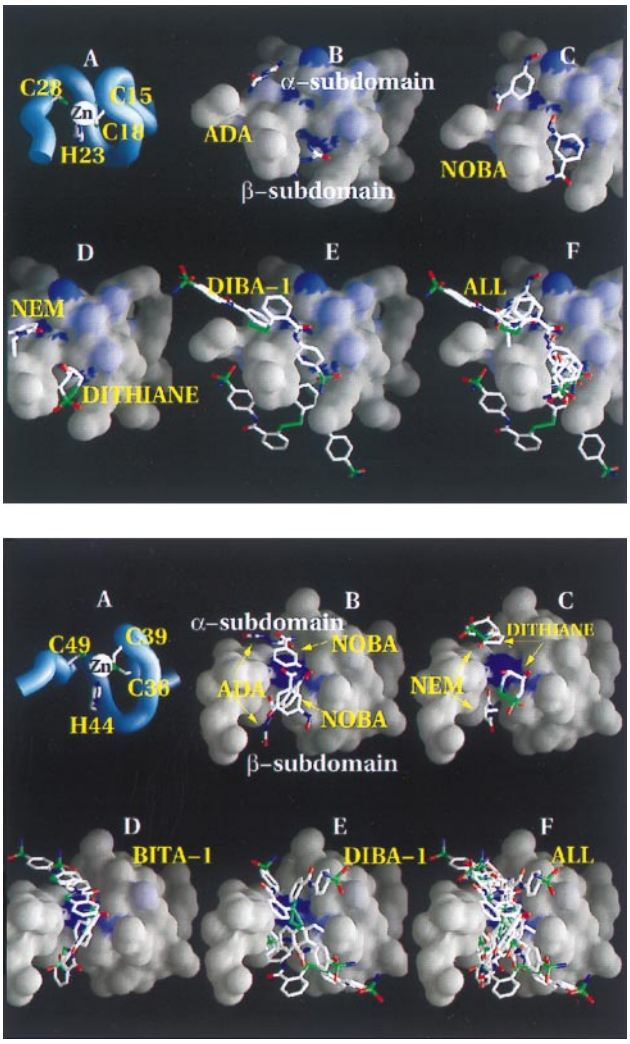

FIG. 4. NCp7–ligand docking arrangements having the best combination of binding affinity and FO overlap. Ligand atom coloring: C white, N blue, O red, and S green. The Connolly surfaces of each finger (0.7 Å probe radius) are colored-coded by $f(r)^-$; dark blue regions indicate the most nucleophilic sites (Cys thiolates) and white the least. (*Upper*), N-terminal finger (finger 1). (*A*) Finger 1 orientation. (*B*) ADA docks. (*C*) NOBA docks. (*D*) NEM and dithiane docks. (*E*) DIBA-1 docks. (*F*) Superposition of all docks. (*Lower*), C-terminal finger (finger 2). (*A*) Finger 2 orientation. (*B*) ADA and NOBA docks. (*C*) NEM and dithiane docks. (*D*) BITA-1 docks. (*E*) DIBA-1 docks. (*F*) Superposition of all docks.

§Assuming $k_{\mathrm{A}} = Z_{\mathrm{A}} \exp(-E_{\mathrm{A}}^{\ddagger}/RT)$, the ratio of reaction rates $k_{\mathrm{A}}/k_{\mathrm{A}'}$, $\mathrm{A} \neq \mathrm{A}'$, for the case $Z_{\mathrm{A}} \simeq Z_{\mathrm{A}'}$ and $E_{\mathrm{A}}^{*} \simeq E_{\mathrm{A}'}^{*}$, has the form $\ln\left(\frac{k_{\mathrm{A}}}{k_{\mathrm{A}'}}\right) \simeq \frac{\zeta}{4RT}\left(\frac{(\chi_{\mathrm{A}}-\chi_{\mathrm{p7}})^2}{\eta_{\mathrm{A}}+\eta_{\mathrm{p7}}} - \frac{(\chi_{\mathrm{A}'}-\chi_{\mathrm{p7}})^2}{\eta_{\mathrm{A}'}+\eta_{\mathrm{p7}}}\right)$, which corresponds to $\Delta\Delta E_{\mathrm{covalent}}$. Temperature-dependence studies are needed to support the assumptions.

Table 2. Fukui nucleophilic indices of the NCp7 zinc finger thiolates

Cysteine sulfur*	$f^-(r)$		
	Structure 1†	Structure 2‡	Structure 3§
Finger 1			
S^{15}	0.37	0.32	0.35
S^{18}	0.29	0.28	0.28
S^{28}	0.34	0.40	0.37
Finger 2			
S^{36}	0.32	0.32	0.23
S^{39}	0.29	0.36	0.28
S^{49}	0.39	0.32	0.49

Based on DFT calculations of the NMR structure (1AAF) (33). Renormalized Fukui sulfur indices, $\Sigma f^-_{S_k} = 1$.

*Cys thiolates labeled by residue.

†Zinc finger domains modeled by the 56-atom [(CCHC)Zn]$^-$ units, Protein Data Bank (PDB) entry no. 1 (reference structure).

‡Same as structure 1, but using PDB entry no. 18 for N-terminal domain (all-atom rms deviation = 0.54 Å), PDB entry no. 11 for C-terminal domain (all-atom rms deviation = 1.13 Å).

§Zinc finger domains modeled by the Lys^{14}-Pro^{31} N-terminal domain (186 atoms) and the Gly^{35}–Thr^{50} C-terminal domain (174 atoms), PDB entry no. 1.

straint $\zeta \equiv 1$, and Gaussian-weighting each pair $(k_A, k_{A'})$ according to their deviation from the least-squares line, yielded $\chi_{p7} = 0.086 \pm 0.015$, $\eta_{p7} = 0.089 \pm 0.057$ ($r^2 = 0.84$, $P = 2 \times 10^{-4}$). In comparison, the average computed values, obtained directly from DFT calculations, were $\chi_{p7} = 0.093 \pm 0.045$, $\eta_{p7} = 0.043 \pm 0.011$. Together, the results provide an estimate of the electronic threshold for reaction by Eq. **3** and strongly suggest that the reactivity of the NCp7 zinc fingers is facilitated by a soft covalent reaction mechanism—i.e., the zinc finger thiolates constitute soft nucleophilic sites. Furthermore, the correlations imply that all ligands were similarly able to gain sufficient proximity to the zinc finger thiolates to initiate reaction. Steric exclusion, or conversely, formation of a stable prereaction complex, would result in substantial inhibition or enhancement of reaction, yielding poor correlations.

Regional Reactivity of the NCp7 Zinc Finger Domains. The Fukui function, $f(r)^-$, probes the nucleophilic regions of the zinc fingers most likely to donate electron density and is displayed in Fig. 4. The Cys thiolates dominated the NCp7 reactivity profile; contributions from the Zn-coordinated His residues were negligible. Given the distorted C_3 symmetry of the local Zn(CCHC) coordination, $f(r)^-$ was considered a more robust measure of local reactivity than ρ_{HOMO}, since molecular orbital theory predicts a triplet of high-lying molecular orbitals derived from mixing of E and A-symmetry HOMOs. Calculations show a close triplet of HOMOs with 80% component due to the Cys sulfurs' $3p$ AOs; their response to electron donation is averaged by $f(r)^-$. The Fukui indices f_k^- (Table 2) of finger 1 were relatively invariant to the local structural variation and domain size modeled; the first and last CCHC thiolates were consistently larger ($S^{15} \sim S^{28} > S^{18}$), but only slightly (10%). In contrast, there was more variation in the f_k^- of finger 2, due to greater structural variation within the set of NMR structures (33). In the largest finger 2 calculation, the index of thiolate S^{49} was significantly larger (20%) than its CCHC counterparts ($S^{49} > S^{36} \sim S^{39}$). Notably, S^{49} does not share electron density via backbone–NH hydrogen bonds (33). Topographically, S^{49} and S^{39} defined a more prominent reactivity surface than found in finger 1 (Fig. 4).

To evaluate thiolate accessibility to electrophilic attack, molecular docking and FO analysis were used. Ligand docking arrangements, having the best overall scores of ΔG_{bind} and FO overlap, are summarized in Table 3 and displayed in Fig. 4 (coordinates available upon request at: www-lecb.ncifcrf.gov/~maynarda/NCp7.html). All chemotypes were generally able to dock with their reactive centers adjacent to zinc finger thiolates, particularly for finger 2. Moreover, orientation of FOs was often complementary. Notably, the computed binding affinities of the best-scoring docks were uniformly weak ($\Delta G_{bind} \geq -2$ kcal/mol). At least for the ligands considered, this observation suggested that electrophilic approach to the Cys thiolates was dictated by steric accessibility rather

Table 3. Summary of NCp7 zinc finger docking arrangements

Ligand (site)	Finger 1 docks			Finger 2 docks		
	Site*	Proximity†	% FO‡	Site*	Proximity†	% FO‡
ADA (N═N)	α	S^{28} 4.8	0.03	α	S^{49} 4.2	0.54
	β	S^{18} 4.3	0.46	β	S^{49} 6.1	0.05
	γ	S^{15} 6.2	0.17	—	—	—
NOBA (*N*═O)	α	S^{28} 6.6	0.02	α	S^{39} 3.3, S^{49} 5.3	0.50
	β	S^{18} 5.5, S^{28} 5.8	0.18	β	S^{39} 3.4, S^{49} 5.2	0.64
	γ	S^{15} 5.3	0.11	β	S^{39} 5.3	0.12
NEM (C═C)	α	S^{18} 6.7, S^{28} 3.9	0.15	α	S^{39} 5.9, S^{49} 3.6	0.76
	γ	S^{15} 6.7	0.02	β	S^{49} 5.3	0.56
Dithiane (*S*—SO_2)	β	S^{18} 5.3	0.34	α	S^{49} 4.3	0.51
	γ	S^{15} 7.6	0.22	β	S^{39} 5.4, S^{49} 5.2	0.79
DIBA-1 (S—S)	α/β	S^{18} 8.2, S^{28} 9.6	0.02	α/β	S^{39} 4.8, S^{49} 5.5	0.33
	β	S^{18} 8.4	0.20	α/β	S^{39} 5.5, S^{49} 5.9	0.14
	γ	S^{15} 8.4	0.14	α/β	S^{39} 6.5, S^{49} 5.4	0.13
BITA-1 (S)	γ	S^{15} 7.8	0.04	α	S^{49} 5.1	0.18
	—	—	—	α	S^{39} 6.5, S^{49} 5.0	0.05
	—	—	—	β	S^{39} 4.3, S^{49} 5.0	0.18
	—	—	—	β	S^{49} 4.6	0.11

Data indicate the docking subdomains, proximity of the ligand electrophilic sites to Cys thiolates, and FO overlap.

*Docking subdomains. Finger 1 (N-terminal) subdomains: $\alpha = N^{17}$, C^{28}–A^{30}, R^{32}; $\beta = N^{17}$, C^{18}, K^{20}, H^{23}, P^{31}, R^{32}; $\gamma = I^{12}$, K^{14}–F^{16}, I^{24}–K^{26}. Finger 2 (C-terminal) subdomains: $\alpha = K^{38}$, C^{39}, K^{41}, M^{46}, C^{49}–E^{51}; $\beta = K^{38}$, K^{41}, H^{44}, D^{48}, C^{49}, E^{51}.

†Cys thiolates (index by residue) in closest proximity to the ligand electrophilic site, relative distance of these sites in Å.

‡Normalized FO overlaps (HOMO-LUMO) expressed as a percentage (100% is complete overlap).

than stable (complex) binding. These results are consistent with the correlations above. Furthermore, NMR analysis (13) and kinetics experiments (14, 36) were unable to detect any significant prereaction ligand binding. The most productive docking domain of each zinc finger had a "saddle-shaped" topography, consisting of α and β subdomains (Table 3) located on either side of the saddle and adjacent to the second and last thiolates of each CCHC motif (Fig. 4). The smaller ligands were able to dock favorably to either subdomain of finger 2, sharing coplanar docking arrangements, while DIBA-1 was capable of utilizing both subdomains. Moreover, the α/β subdomains of finger 1 did not equally allow docking of all chemotypes. BITA-1 was able to dock favorably only to a third, γ, subdomain of finger 1, located on the opposite side of the saddle (Table 3), recently identified as a specific RNA (guanosine) binding site (37). Although all chemotypes docked favorably with the finger 1 γ subdomain, most were far removed from S^{15} (the nearest thiolate). Table 3 can be used to compare reactivities of fingers 1 and 2. In all cases, the closest reactive ligand approaches to the Cys thiolates occurred at finger 2, particularly for DIBA-1 and BITA-1. As a consequence, larger FO overlaps were observed for finger 2 and S^{49} tended to be the thiolate in closest proximity to the ligand reactive centers. On the basis of both reactive accessibility and electronic response, the results suggest that S^{49} is the NCp7 site most labile to electrophilic attack. These findings are consistent with observations that Cys^{49} preferentially reacted with 2,2′-dithiopyridine (15), and finger 2 was 7-fold more reactive than finger 1 towards DIBA-2 and BITA-1 (14, 36).

Further pursuit of the methods in this study may help identify new compounds having sufficient electronic and structural properties that favor reaction with NCp7. The results suggest that inclusion of softness in the criteria of electrophile design may improve the efficacy of therapeutics. Experimental comparison of the reactivities of the NCp7 zinc fingers, versus a "structureless" soft nucleophile, could also provide a valuable empirical free-energy relationship for resolving the steric component of reactivity. Notably, significant differences were found between the steric accessibility of fingers 1 and 2, particularly for the larger BITA and DIBA ligands, even though both zinc fingers share the same motif. A recent study (16) also demonstrated that ADA, dithiane, and DIBA-1 selectively inhibited NCp7 without affecting other cellular zinc finger proteins. Thus, the electronic and structural features of some conserved zinc fingers may prove sufficiently unique to allow specific therapeutic targeting.

We thank the staff of the Frederick Biomedical Supercomputing Center for resources and technical support. Special thanks are given to Yongsheng Song and John Inman for reagents, and Anders Wallqvist, Tom Mallouk, Weitao Yang, and Igor Topol for helpful discussions. A.T.M. is grateful for funding by a National Institutes of Health Postdoctoral Fellowship. This research was sponsored by the National Cancer Institute.

1. Weiss, R., Teich, N., Varmus, H. & Coffin, J. (1984) *RNA Tumor Viruses* (Cold Spring Harbor Lab. Press, Plainview, NY).
2. Cohen, J. (1997) *Science* **277,** 32–33.
3. Darlix, J., Lapada-Tapolsky, M., de Rocquigny, H. & Roques, B. (1995) *J. Mol. Biol.* **254,** 523–537.
4. Berg, J. M. (1986) *Science* **232,** 485–487.
5. Henderson, L., Copeland, T., Sowder, R., Smythers, G. & Oroszlan, S. (1981) *J. Biol. Chem.* **256,** 8400–8406.
6. South, T., Blake, P., Sowder, R., Arthur, L., Henderson, L. & Summers, M. (1990) *Biochemistry* **29,** 7786–7789.
7. Rice, W. G., Schaeffer, C., Harten, B., Villinger, F., South, T., Summers, M., Henderson, L., Bess, J., Arthur, L., McDougal, J., *et al.* (1993) *Nature (London)* **361,** 473–475.
8. Rice, W. G., Turpin, J., Schaffer, C., Graham, L., Clanton, D., Buckheit, R., Zaharevitz, D., Summers, M., Wallqvist, A. & Covell, D. (1996) *J. Med. Chem.* **39,** 3606–3616.
9. Rice, W. G., Baker, D., Scheaffer, C., Graham, L., Bu, M., Terpening, S., Clanton, D., Schultz, R., Bader, J., Buckheit, R., *et al.* (1997) *Antimicrob. Agents Chemother.* **41,** 419–426.
10. Rice, W. G., Turpin, J., Huang, M., Clanton, D., Buckheit, R., Covell, D., Wallqvist, A., McDonnell, N., Guzman, R. D., Summers, M., *et al.* (1997) *Nat. Med.* **3,** 341–345.
11. Turpin, J. A., Terpening, S., Schaeffer, C., Yu, G., Glover, C., Felsted, R, Sausville, E. & Rice, W. (1996) *J. Virol.* **70,** 6180–6189.
12. Turpin, J. A., Schaeffer, C., Terpening, S., Graham, L., Bu, M. & Rice, W. (1997) *Antiviral. Chem. Chemother.* **8,** 60–69.
13. Loo, J. A., Holler, T., Sanchez, J., Gogliotti, R., Maloney, L. & Reily, M. (1996) *J. Med. Chem.* **39,** 4313–4320.
14. Tummino, P. J., Scholten, J., Harvey, P., Holler, T., Maloney, L., Gogliotti, R., Domagala, J. & Hupe, D. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 969–973.
15. Hathout, Y., Fabris, D., Han, M., Sowder, R., Henderson, L. & Fenselau, C. (1997) *Drug Metab. Dispos.* **24,** 1395–1400.
16. Huang, M., Maynard, A., Turpin, J., Graham, L., Janini, G., Covell, D. & Rice, W. G. (1998) *J. Med. Chem.* **41,** 1371–1381.
17. Pearson, R. G. (1990) *Coord. Chem. Rev.* **100,** 403–425.
18. Parr, R. G. & Pearson, R. G. (1983) *J. Am. Chem. Soc.* **105,** 7512–7516.
19. Parr, R. G. & Yang, W. (1989) *Density-Functional Theory of Atoms and Molecules* (Oxford Univ. Press, Oxford).
20. Kohn, W., Becke, A. & Parr, R. G. (1996) *J. Phys. Chem.* **100,** 12974–12980.
21. Pearson, R. G. (1988) *Inorg. Chem.* **27,** 734–740.
22. Berg, J. M. & Lippard, S. J. (1994) *Principles of Bioorganic Chemistry* (University Science Books, Mill Valley, CA).
23. Leis, J., Pena, M. & Rios, A. (1995) *J. Chem. Soc. Perkin Trans. 2* **116,** 587–593.
24. Vandevelde, M., Witvrouw, M., Schmit, J., Sprecher, S., DeClercq, E. & Tassignon, J. (1996) *AIDS Res. Hum. Retroviruses* **12,** 567–568.
25. Sanderson, R. T. (1983) *Polar Covalence* (Academic, New York).
26. Parr, R. G. & Yang, W. (1984) *J. Am. Chem. Soc.* **106,** 4049–4050.
27. Lee, C., Yang, W. & Parr, R. G. (1988) *J. Mol. Struct.* **163,** 305–313.
28. Damoun, S., Van de Woude, G., Mendez, F. & Geerlings, P. (1997) *Phys. Chem. A* **101,** 886–893.
29. Delley, B. (1990) *J. Chem. Phys.* **92,** 508–517.
30. Vosko, S., Wilk, L. & Nusair, M. (1980) *Can. J. Phys.* **58,** 1200–1211.
31. Becke, A. D. (1988) *J. Chem. Phys.* **88,** 2547–2553.
32. Perdew, J. P. & Wang, Y. (1992) *Phys. Rev. B* **45,** 13244–13249.
33. Summers, M. F., Henderson, L., Chance, M., Bess, J., South, T., Blake, P., Perez-Alvarado, G. & Sowder, R. (1992) *Protein Sci.* **1,** 563–574.
34. Wallqvist, A. & Covell, D. G. (1996) *Proteins* **25,** 403–419.
35. Vuilleumier, C., Maechling-Strasser, C., Gerard, D. & Mely, Y. (1997) *Anal. Biochem.* **244,** 183–185.
36. Tummino, P. J., Harvey, P., McQuade, T., Domagala, J., Gogliotti, R., Sanchez, J., Song, Y. & Hupe, D. (1997) *Antimicrob. Agents Chemother.* **41,** 394–400.
37. DeGuzman, R. N., Wu, Z., Stalling, C., L. Pappalarado, P. & Summers, M. F. (1998) *Science* **279,** 384–388.