## Mechanisms of Directed Attention in the Human Extrastriate Cortex as Revealed by Functional MRI

Sabine Kastner,\* Peter De Weerd, Robert Desimone, Leslie G. Ungerleider

A typical scene contains many different objects, but the capacity of the visual system to process multiple stimuli at a given time is limited. Thus, attentional mechanisms are required to select relevant objects from among the many objects competing for visual processing. Evidence from functional magnetic resonance imaging (MRI) in humans showed that when multiple stimuli are present simultaneously in the visual field, their cortical representations within the object recognition pathway interact in a competitive, suppressive fashion. Directing attention to one of the stimuli counteracts the suppressive influence of nearby stimuli. This mechanism may serve to filter out irrelevant information in cluttered visual scenes.

The human visual system is usually confronted with cluttered scenes consisting of many different objects, which cannot all be processed simultaneously. Only a limited amount of what we see reaches consciousness and becomes stored in memory, which indicates that there is limited processing capacity within the visual system and that multiple object representations are in competition for access to this limitedcapacity system (1). One way to resolve the competition is through spatially directed attention. If one attends, for example, to a specific location in a cluttered scene, information processing is greatly facilitated at that location, while interfering information from objects at nearby locations is efficiently filtered out. This suggests that processing is biased in favor of the attended location (2).

Results from single-cell recordings in extrastriate cortical areas in the ventral object vision pathway of monkeys are consistent with these ideas (3). Evidence for competition is provided by the finding that the response to an otherwise optimal stimulus presented within a neuron's receptive field is often reduced when a second stimulus is presented simultaneously at a different location within the same receptive field. Hence, multiple stimuli are not processed independently from each other but rather interact competitively in a mutually suppressive fashion. This competition can be biased in favor of one of

the stimuli by spatially directed attention. If an animal directs its attention to one of the competing stimuli within the receptive field, the responses are as large as those to the stimulus presented alone. These results suggest that spatially directed attention to a visual stimulus cancels out the suppressive influence of nearby stimuli, thereby enhancing information processing at the attended location. If so, this could be a mechanism to filter out unwanted information in cluttered visual scenes.

We used functional magnetic resonance imaging (fMRI) in humans to test for the presence of suppressive interactions among stimuli presented simultaneously within the visual field in the absence of directed attention (experiment 1) and to investigate the influence of spatially directed attention on these suppressive interactions (experiment 2). The design for experiment 1 is presented in Fig. 1. Complex visual images were shown in randomized order in four nearby locations within the right upper quadrant under two presentation conditions: sequential and simultaneous (Fig. 1, A and B). In the sequential condition (SEQ), each of the stimuli was shown alone in one of the four locations. In the simultaneous condition (SIM), the stimuli appeared together in all four locations. Integrated over time, the physical stimulation parameters in each of the four locations were identical under the two conditions. However, suppressive interactions among stimuli could take place in the simultaneous but not in the sequential condition. Thus, on the basis of the results from monkey physiology (4), we hypothesized that the fMRI signals would be smaller during the simultaneous than during the sequential presentations because of the

mutual suppression induced by competitively interacting stimuli.

Functional MRI scans were obtained from eight people, and data were analyzed by means of multiple regression (5). Sequential and simultaneous conditions were presented in blocks of 18 s each, interleaved with equally long blank periods in the sequence SEQ-SIM-SIM-SEQ. The participant's task was to count T's or L's at the fixation point throughout the scan, which fully engaged the participant's attention at fixation and not at the peripherally presented stimuli (6).

The visual areas that were consistently activated in all participants in the ventral striate and extrastriate cortex during visual stimulation as compared to blank periods were in the calcarine sulcus [Brodmann area (BA) 17], the lingual gyrus (BA 18), and the fusiform gyrus (BA 19 and 37) of the left hemisphere, as illustrated for a single participant in Fig. 2A. Also shown is the assignment of activated voxels to areas V1 to TEO on the basis of meridian mapping (7), which was performed in a separate scan session for each participant (mean Talairach coordinates across all participants were as follows: V1: x = -3, y = -81, z = +8; V2: -9, -78, -10; V4: -19, -74, -14; TEO: -27, -59, -14). As predicted by our hypothesis that stimuli presented together interact in a mutually suppressive way, simultaneous presentations evoked weaker responses than sequential presentations, as shown by the averaged time series of fMRI signals (Fig. 2B, left panel) and by the mean signal differences (Fig. 3A), which were significant in all areas [repeated measures analysis of variance (ANOVA); P < 0.01 for V1, V2, and TEO; P < 0.001 for V4]. The difference in activations between sequential and simultaneous presentations increased from V1 to V4 and TEO (Fig. 3A) [interaction of cortical area and presentation condition: F(3, 15) = 25.1, P < 0.001]; this effect is also reflected in the sensory suppression index (Fig. 3C) [SSI =  $(R_{SEQ} - R_{SIM})$ /  $(R_{\rm SEQ} + R_{\rm SIM}); R =$ averaged responses of the peak MRI intensities obtained during visual presentation blocks for a given condition]. The increase in the magnitude of the suppression index across visual areas suggests that the suppressive interactions were scaled to the increase in receptive field size across these areas. Because of their small receptive fields, individual neurons in V1 and V2 would be capable of processing information only from a very limited portion of our  $4^{\circ} \times 4^{\circ}$  display, resulting in minimal interaction effects between stimuli; whereas neurons in V4 and TEO, with their larger receptive fields, would process information from all four stimuli in the display, resulting in significantly greater suppressive interaction effects (SSI: V1/V2 versus V4/TEO, P <0.0001). In further support of this idea, in-

S. Kastner, P. De Weerd, L. G. Ungerleider, Laboratory of Brain & Cognition, National Institute of Mental Health (NIMH), National Institutes of Health (NIH), Building 49, Room 1B80, Bethesda, MD 20892–4415, USA. R. Desimone, Laboratory of Neuropsychology, NIMH, NIH, Bethesda, MD 20892, USA.

<sup>\*</sup>To whom correspondence should be addressed. E-mail: sabine@ln.nimh.nih.gov

creasing the separation between stimuli decreased the suppressive interactions (8).

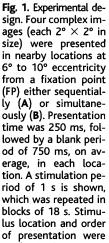
In both the sequential and simultaneous conditions, the stimulus presentation rate at any one of the four locations was 1 Hz. However, across the visual field the overall presentation rate in the two conditions was different. To rule out the possibility that the differential responses evoked by the two presentation conditions reflected differences in overall stimulus presentation rate, we sought to demonstrate suppressive interactions directly in a control experiment, in which the presentation rate was held constant. In this experiment, we presented one of the stimuli close to the horizontal meridian in the upper visual field in the absence and in the presence of three other stimuli presented nearby in the lower visual field (Fig. 4); under both conditions, the stimuli were presented at a rate of 1 Hz (9). Because extrafoveal upper and lower visual field representations within early extrastriate areas are located in spatially separated regions, nearby stimuli placed on opposite sides of the horizontal meridian may competitively interact but evoke activations that are separable in the cortex. As shown in Fig. 4 for a single participant, the response evoked in V4's upper field by the single stimulus was significantly greater than the response evoked by the same stimulus presented together with the three stimuli in the lower visual field. The averaged signal changes for all participants tested (n = 3)were significantly different in the two conditions in V4's upper field (paired t test, P <0.01) (10). This finding supports the idea of suppressive interactions among the stimuli and cannot be explained by stimulus presentation rate.

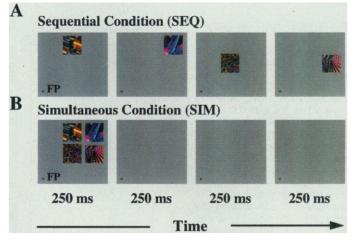
To study the influence of spatially directed attention on suppressive interactions between stimuli, five of the eight participants were tested in experiment 2. This experiment employed a factorial design with two main factors—presentation condition (sequential versus simultaneous) and directed attention condition (unattended versus attended). During each scan, the four blocks of visual stimulation (SEQ-SIM-SIM-SEQ) were tested in an unattended and an attended condition, with the order of the two conditions being counterbalanced across scans (11). In the unattended condition, attention was directed away from the location of the stimuli by having participants count T's or L's at the fixation point, just as in experiment 1. In the attended condition, participants were instructed to covertly attend to the location of the stimulus in the array that was closest to the fixation point and to count the occurrences of a particular target stimulus at that location (12). The target stimulus was indicated by its presentation before each scan. We hypothesized that spatially directing attention to

stimuli at one location in the four-element array would reduce the suppressive effects of the surrounding stimuli on the target stimulus in the simultaneous condition (13). Hence, we predicted that attention would enhance the responses to simultaneously presented stimuli more strongly than to sequentially presented stimuli.

In accordance with our hypothesis, the averaged fMRI signal with attention in V4 and TEO increased by 0.84 and 0.62%, respectively, to simultaneously presented stimuli but only by 0.48 and 0.34%, respectively, to sequentially presented stimuli (Fig. 3B). The interaction between the attention and

presentation factors was significant in areas V4 (Fig. 2B, blue shaded blocks) and TEO [repeated measures ANOVA; V4: F(1, 4) =11.2, P < 0.05; TEO: F(1, 4) = 8.5, P <0.05] but just failed to reach significance in V2 [F(1, 4) = 7.5, P = 0.052]. Thus, the suppressive interactions were partially canceled out by attention. This is also demonstrated by the reduced SSIs in the attended as compared to the unattended condition shown in Fig. 3D. This figure also shows that the magnitude of the attentional effect scaled with the magnitude of the suppressive interactions between stimuli, with the strongest reduction of suppression occurring in V4 and





randomized. New images were chosen out of a pool of 100 for different runs.

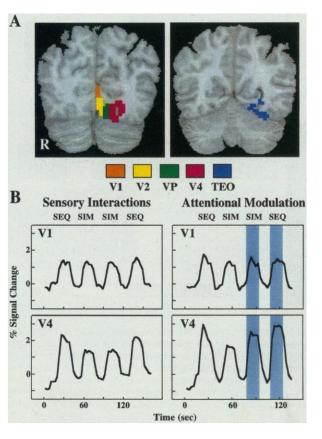


Fig. 2. (A) Brain areas activated by the complex images as compared to blank presentations. Coronal slices of a single participant at a distance of 25 mm (left) and 40 mm (right) from the posterior pole. Activated voxels were assigned to areas V1, V2, VP, V4, and putative TEO by meridian mapping (7). R indicates right hemisphere. (B) Time series of fMRI signals in V1 and V4 in experiment 1 (left) and experiment 2 (right), averaged over all participants. In experiment 1, sequentially presented stimuli evoked stronger activations than did simultaneously presented stimuli. This effect was much stronger in V4 than in V1 and was replicated in the unattended condition of experiment 2 (unshaded time series). Spatially directed attention (blue shading) increased responses to simultaneously presented stimuli to a larger extent than to sequentially presented ones in V4. Presentation blocks were 18 s in experiment 1 and 15 s in experiment 2.

TEO. The results therefore support the second hypothesis that spatially directed attention enhances processing of stimuli in the attended location by counteracting suppression induced by nearby stimuli.

We also found a general increase in activity, affecting the response under both sequential and simultaneous conditions [repeated measures ANOVA; main attentional effect: F(1, 4) = 17.2, P < 0.05] with a significant interaction between cortical area and attentional effect [F(3, 12) = 6.2, P < 0.01] (Fig. 2B, blue shaded blocks; Fig. 3B). The effect of attention was significant in areas V2 (P < 0.05), V4 (P < 0.01), and TEO (P < 0.05) but not in V1 (P = 0.83) (14). These results

are consistent with single-cell, event-related potential and imaging studies that have found enhanced responses or increased baseline activity in the ventral extrastriate cortex in response to stimuli presented at attended locations (3, 4, 15).

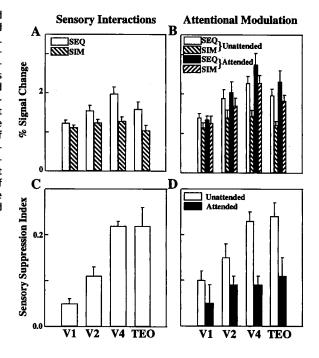
Our results indicate that, in the absence of directed attention, multiple stimuli in the visual field interact with each other in a mutually suppressive way, as demonstrated by the reduced fMRI signals to simultaneously presented stimuli as compared to sequentially presented ones. Spatially directed attention reduces these interactions by partially canceling out their suppressive effects, as demonstrated by significantly greater effects of at-

tention on the fMRI signal evoked by simultaneously presented stimuli as compared to that evoked by sequentially presented ones. Both the sensory interactions and attentional effects scale with the sizes of the neuronal receptive fields along the ventral object vision pathway. Modulation of suppression at several extrastriate stages may therefore be a mechanism by which attention filters out unwanted information.

## References and Notes

- J. Duncan, Psychol. Rev. 87, 272 (1980); \_\_\_\_\_\_ and G. Humphreys, ibid. 96, 433 (1989); R. Desimone and J. Duncan, Annu. Rev. Neurosci. 18, 193 (1995); G. Rees, C. D. Frith, N. Lavie, Science 278, 1616 (1997).
- M. I. Posner, Quart. J. Exp. Psychol. 32, 3 (1980); A. Treisman, Comput. Vision Graphics Image Process. 31, 156 (1985); J. Driver and G. C. Baylis, J. Exp. Psychology 15, 448 (1989).
- J. Reynolds, L. Chelazzi, S. J. Luck, R. Desimone, Soc. Neurosci. Abstr. 21, 1759 (1995); J. Moran and R. Desimone, Science 229, 782 (1985); S. Treue and J. H. R. Maunsell, Nature 382, 539 (1996); C. E. Connor, D. C. Preddie, J. L. Gallant, D. C. Van Essen, J. Neurosci. 17, 3201 (1997); E. K. Miller, P. M. Gochin, C. G. Gross, Brain Res. 616, 25 (1993).
- 4. J. S. J. Luck, L. Chelazzi, S. A. Hillyard, and R. Desimone [J. Neurophysiol. 77, 24 (1997)] reported that both suppressive interactions and attentional effects were larger in V4 when competing stimuli were presented simultaneously than when presented sequentially at different locations. It is not known whether the suppressive effects are due to inhibition or reduction of excitation.
- Fourteen contiguous, coronal, 5-mm-thick slices were acquired in 12 to 16 series of 60 images each, starting from the posterior pole (in-plane resolution, 2.5 mm imes 2.5 mm). Gradient echo, echo planar imaging was used [reception time (TR) = 3 s, echo time (TE) = 40 ms, flip angle = 90°] on a 1.5 Tesla GE magnet. Functional images were coaligned with a high-resolution anatomical scan taken in the same session (three-dimensional spoiled gradient echo sequence; TR = 15 ms; TE = 7 ms; flip angle = 30°; matrix, 256 × 256 voxels). Activations were identified by means of multiple regression analysis of the time series of MRI intensities in every voxel and two regressors of interest [K. J. Friston et al., Neuroimage 2, 45 (1995)], reflecting contrasts between (i) visual stimulation versus blank periods and (ii) sequential versus simultaneous presentations. Additional regressors were used to factor out variance due to between-run changes in mean intensity and within-run linear changes. The statistical significance (P < 0.05) of activated regions was assessed by an analysis based on the spatial extent of each region [K. Friston, K. J. Worsley, R. S. J. Frackowiak, J. C. Mazziotta, A. C. Evans, Hum. Brain Mapp. 1, 210 (1994)]. The fMRI time series, averaged over all activated voxels in a given region during visual stimulation versus blank presentations (thresholded at a Z score of 3.7) and over runs for each participant, are presented as group data. In experiment 2, time series analysis was performed only on those voxels activated in both the unattended and attended conditions. Statistical significance was assessed with repeated measures ANO-VAs on the six (experiment 1) or five (experiment 2) peak intensities of the fMRI signal. For each subject, statistical maps and structural images were transformed into Talairach space [J. Talairach and P. Tournoux, Co-Planar Stereotactic Atlas of the Human Brain (Thieme, New York, 1988)] using the template from SPM96. Participants (four men, 22 to 34 years old) gave written informed consent for participation.
- 6. T's and L's (0.6° in size) were presented for 250 ms in random order and in different orientations at 4 Hz. The T/L task had a high attentional load [see also (12)], in order to ensure proper fixation and to prevent participants from covertly attending to the peripheral stimuli. Performance measured outside the

Fig. 3. Mean signal changes and SSIs in areas V1, V2, V4, and TEO, averaged over participants. Results are shown for experiment 1 (A and C) and experiment 2 (B and D). Vertical bars indicate SEM. SSIs increased from V1 to V4 and TEO in experiment 1, which suggests that the effects were scaled to the increasing receptive field sizes of neurons in these areas. This finding was replicated in the unattended condition of experiment 2. In the attended condition of experiment 2, SSIs showed the strongest reduction in V4 and TEO.



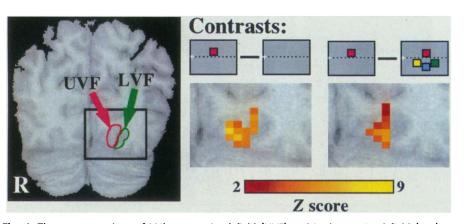


Fig. 4. The representations of V4's upper visual field (UVF) and its lower visual field (LVF) are located medially and laterally in separated but neighboring locations on the fusiform gyrus (left panel). The activity evoked by a single stimulus ( $2^{\circ} \times 2^{\circ}$ ) presented at  $8^{\circ}$  eccentricity just above the horizontal meridian, as compared to blank presentations, was confined to V4's UVF representation (middle panel). As demonstrated in this participant, more activity was evoked in V4's UVF when the stimulus was presented alone than when it was shown together with three stimuli in the LVF just below the horizontal meridian (right panel). In all presentation conditions, stimuli were presented for 250 ms at 1 Hz.

## REPORTS

- scanner (75% correct on average) did not differ during blank, sequential, or simultaneous presentation periods [F(2, 143) = 1.60, P = 0.21]. Hence, neither presentation condition interfered with the T/L task, indicating that this task provided sufficient attentional load to preclude exogenous attentional cueing.
- 7. The borders of retinotopic areas in the ventral extrastriate cortex of humans and monkeys [R. Gattass, C. G. Gross, J. H. Sandell, J. Comp. Neurol. 21, 519 (1981); R. Gattass, A. P. Sousa, C. G. Gross, J. Neurosci. 8, 1831 (1988); M. I. Sereno et al., Science 268, 889 (1995); R. B. H. Tootell et al., J. Neurosci. 15, 3215 (1995); E. A. DeYoe et al., Proc. Natl. Acad. Sci. U.S.A. 93, 2382 (1996)] are formed by the representations of either the vertical (V1/V2 or VP/V4) or the horizontal (V2/VP) meridians. Meridians were mapped with color- and luminance-contrast checkered stimuli. In five of eight participants, it was difficult to determine the extent of VP, because the representations of the V2/VP and the VP/V4 border were abutting or overlapping [see S. Shipp, J. D. G. Watson, R. S. J. Frackowiak, S. Zeki, Neuroimage 2, 125 (1995)]. We will, therefore, refer to the area between the V1/V2 border and the VP/V4 border as "V2", although it likely contains parts of VP. The presumptive lower field representation of V4 was determined with the complex images presented to the lower field and was found to be located adjacent and lateral to V4's upper field representation on the fusiform gyrus [D. J. McKeefry and S. Zeki, Brain 120, 2229 (1997)]. The region we have termed V4 may include all or part of the region termed V8 by Hadjikhani et al. [Nature Neurosci. 1, 235 (1998)]. In the region located anterior to V4 (and also V8), the spatial segregation of upper and lower field representations was no longer seen, suggesting that this area was different from V4. Because area TEO is located just anterior to V4 in the monkey [D. Boussaoud, R. Desimone, L. G. Ungerleider, J. Comp. Neurol. 306, 554 (1991)], we will refer to this similarly located area as putative human TEO.
- 8. In a separate experiment, four stimuli (each  $0.5^{\circ} \times 0.5^{\circ}$  in size) were presented 6° apart from each other in the right upper quadrant. The prediction was that increasing the spatial separation between stimuli would strongly reduce suppressive interactions in areas with small (V2) and intermediate (V4) receptive fields but not in areas with large receptive fields (TEO) extending over a quadrant. Results from three participants showed that the interactions were indeed abolished in V2, were strongly reduced in V4, but were still present in TEO.
- 9. Three of the eight participants saw complex stimuli at 1 Hz in the following presentation configurations: one stimulus presented to the upper visual field, three presented to the lower visual field, or all four presented together. Participants performed the T/L task at fixation throughout the scan. All other presentation parameters were as in experiment 1.
- 10. The averaged signal changes in V4's upper field were 1.04% evoked by the single stimulus, 0.83% evoked by the four stimuli, and 0.52% evoked by the three stimuli in the lower field (due to signal spread into the upper field). Because of this spread, the actual suppression effect might be much larger than that reflected in the difference in responses to the single stimulus and to the four stimuli. The response differences were not significant in V1 and V2. Thus, with this experimental design, suppressive interactions could only be demonstrated in areas with sufficiently large receptive fields.
- 11. All four stimuli, including the stimulus selected to be the target, were randomly presented in all four locations in blocks of 15 s. The blocks with directed attention to the stimulus display were indicated by a marker presented close to the fixation point 1 s before the block started. In pilot experiments, we found that the attentional effect during the first attended block in a sequence was always stronger than in other attentional blocks within a run. To attenuate this attentional "onset" effect, each run started with a block of attended presentations that was discarded from analysis.
- 12. Before being scanned, participants received training

- in the directed attention task and fixation was monitored. During the directed attention task, targets were identified correctly at rates of 86 and 93%, respectively, in the sequential and simultaneous presentation conditions. The attentional load of the T/L task and the directed attention task was assessed by having participants perform them simultaneously in tests conducted outside the scanner. Both tasks interfered with each other when performed simultaneously. Performance in the directed attention task dropped significantly [F(1, 192) = 130.92, P <0.0001] from 86 to 45% and from 93 to 49%, respectively, in the sequential and simultaneous conditions. Likewise, performance in the T/L task dropped significantly [F(1, 191) = 66.76, P < 0.0001]when participants were required to simultaneously identify targets at the target location. Thus, both tasks had a high attentional load. Participants rarely identified target stimuli in locations other than the attended location.
- 13. Because the cortical activations from the attended and unattended stimuli could not be separated, any increase in response to the attended stimulus might, in principle, be counterbalanced by a decrease in response to an unattended one, working against our hypothesis. However, the attended stimulus was located closest to the fovea and thus would dominate the response to the array because of the cortical

- magnification factor. Further, single-cell studies have shown that attention to a stimulus filters out the suppressive influence of nearby stimuli very effectively, but it has a smaller suppressive effect on the response to unattended ones (R. Desimone, unpublished observations).
- 14. Cortical volumes activated in the unattended condition were 394 mm³ in V1, 400 mm³ in V2, 1600 mm³ in V4, and 1156 mm³ in TEO, averaged over participants. In the attended condition, brain volumes increased significantly in V4 and TEO but not in V1 and V2 [V4: 78  $\pm$  16% (mean  $\pm$  SEM); TEO: 120  $\pm$  36%; ANOVA, main attentional effect: F(1, 64) = 14.2, P < 0.001; cortical area and attentional effect: F(3, 64) = 2.82, P < 0.05].
- H. J. Heinze et al., Nature 372, 543 (1994); G. R. Mangun, Psychophysiology 32, 4 (1995); G. Rees, R. Frackowiak, C. Frith, Science 275, 835 (1997); R. Vandenberghe et al., J. Neurosci. 17, 3739 (1997).
- 16. We thank J. M. Maisog, M. I. Elizondo, and M. A. Georgopoulos for help with data analysis; P. Jezzard for help with scanning; and J. V. Haxby, B. Jagadeesh, A. Martin, J. Reynolds, and U. Ziemann for valuable discussions. S.K. was supported by the Deutsche Forschungsgemeinschaft.

11 May 1998; accepted 4 August 1998

## A Structural Basis for Recognition of A·T and T·A Base Pairs in the Minor Groove of B-DNA

Clara L. Kielkopf, Sarah White, Jason W. Szewczyk, James M. Turner, Eldon E. Baird, Peter B. Dervan,\*
Douglas C. Rees\*

Polyamide dimers containing three types of aromatic rings—pyrrole, imidazole, and hydroxypyrrole—afford a small-molecule recognition code that discriminates among all four Watson-Crick base pairs in the minor groove. The crystal structure of a specific polyamide dimer-DNA complex establishes the structural basis for distinguishing T·A from A·T base pairs. Specificity for the T·A base pair is achieved by means of distinct hydrogen bonds between pairs of substituted pyrroles on the ligand and the O2 of thymine and N3 of adenine. In addition, shape-selective recognition of an asymmetric cleft between the thymine-O2 and the adenine-C2 was observed. Although hitherto similarities among the base pairs in the minor groove have been emphasized, the structure illustrates differences that allow specific minor groove recognition.

Before the first structure of a molecule bound to DNA had been determined, specific recognition of double helical B-form DNA was

C. L. Kielkopf, Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA. S. White, J. W. Szewczyk, J. M. Turner, E. E. Baird, P. B. Dervan, Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, USA. D. C. Rees, Howard Hughes Medical Institute and Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, USA.

\*To whom correspondence should be addressed. E-mail: dervan@its.caltech.edu (P.B.D.); dcrees@its.caltech.edu (D.C.R.)

predicted to occur primarily in the major, rather than the minor, groove (1). This proposal was based on the observation that for A,T base pairs, the hydrogen bond acceptors at N3 of adenine and O2 of thymine are similarly placed and lack any prominent distinguishing feature (1) (Fig. 1). Subsequent structures of DNA binding domains cocrystallized with DNA supported this idea, because most of the specific contacts were made with the major groove (2). The principle that "the major groove is a better candidate for sequence-specific recognition than the minor groove" (3) continues to provide the basis for strategies to decipher rules for