The Chemistry of John Dalton’s Color Blindness

David M. Hunt, Kanwaljit S. Dulai, James K. Bowmaker, John D. Mollon

The Chemistry of John Dalton’s Color Blindness

David M. Hunt,* Kanwaljit S. Dulai, James K. Bowmaker, John D. Mollon

John Dalton described his own color blindness in 1794. In common with his brother, he confused scarlet with green and pink with blue. Dalton supposed that his vitreous humor was tinted blue, selectively absorbing longer wavelengths. He instructed that his eyes should be examined after his death, but the examination revealed that the humors were perfectly clear. In experiments presented here, DNA extracted from his preserved eye tissue showed that Dalton was a deuteranope, lacking the middlewave photopigment of the retina. This diagnosis is shown to be compatible with the historical record of his phenotype, although it contradicts Thomas Young’s belief that Dalton was a protanope.

Two centuries ago the chemist John Dalton analyzed his own color blindness in his first lecture to the Manchester Literary and Philosophical Society (1), and “daltonism” has since become the name for the condition in many languages (2). Dalton judged red sealing wax to be a good match for the outer face of a laurel leaf, and a crimson ribbon matched the color that others called “mud” (1, 3). In the solar spectrum he saw only two main hues, one of which corresponded to the normal observer’s red, orange, yellow, and green, whereas the second corresponded to blue and violet. He was particularly surprised to observe that the pink flowers of a cranesbill (Geranium zonale) (4), which appeared “sky-blue” to him by daylight, looked “very near yellow, but with a tincture of red” (3) by candlelight (Fig. 1). Of his immediate acquaintances, only his brother shared his astonishment at this failure of color constancy (3).

In an explanation of his color deficiency, Dalton proposed that the vitreous humor of his eye was tinted blue, selectively absorbing longer wavelengths. He instructed that after his death his eyes should be dissected to confirm his hypothesis. He died at age 78 on 27 July 1844, and on the following day an autopsy was done by his medical attendant, Joseph Ransome (5, 6). Ransome collected the humors of one eye into watch glasses and found them to be “perfectly pellucid,” the lens itself exhibiting the yellowness expected in someone of Dalton’s age (6). He shrewdly left the second eye almost intact, slicing off the posterior pole and noting that scarlet and green objects were not distorted in color when seen through the eye. Thus, Ransome found no support for Dalton’s hypothesis that color blindness was due to a preretinal filter. Ransome did not discard the eyes but stored them only in air, and fragments of them have survived to this day (Fig. 2). Originality in the possession of Dalton Hall, the eyes passed into the keeping of the Manchester Literary and Philosophical Society (7) who gave us permission to take small samples for a reexamination of Dalton’s color blindness by DNA analysis.

At the time of Dalton’s death, the most popular alternative to his own hypothesis was that daltonism arose from a cerebral defect. In his autopsy report, Ransome recorded (6) a “deficient development” of the phrenological organ of color (which corresponded to one of the convolutions of the frontal lobe). Today we know that hereditary color blindness seldom arises from either a preretinal filter or a central defect, but rather is due to the absence or the alteration of one of the photosensitive pigments of the retina. Interestingly, an explanation of this kind had already been proposed by 1794, for although we owe to Dalton the first systematic analysis of the confusions that occur in color blindness, the actual existence of the condition was well enough known for King George III to have raised it in conversation with Fanny Burney at court in 1785 (8). A little earlier a German general science magazine reported the theory advanced by a shadowy figure called Giros von Gentilly (9) who held that there were three types of “molecule” or “membrane” in the retina, corresponding to three kinds of light. Color blindness arose when one or two of the three types of molecule was either paralyzed or constitutively overactive (10).

Thomas Young, author of the wave theory of light and the decipherer of the Rosetta stone, may have heard of von Gentilly’s ideas while a medical student in Goettingen. Young combined a three-receptor theory of vision with a realization that the physical correlate of hue is a continuous variable. To explain Dalton’s color blindness, Young postulated “the absence or paralysis of some fibers of the retina, which are calculated to perceive red” (11). In supposing that it was the red-sensitive fibers that were affected, Young was critically influenced by one phrase in Dalton’s description of the solar spectrum: “that part of the image which others call red, appears to me little more than a shade, or defect of light” (1).

After much debate, the three-receptor theory of von Gentilly and of Young is today generally accepted. We now know that daylight vision depends on three types of photopigments, segregated into three classes of cone cell, with peak wavelength sensitivities (λmax) near 420 nm (shortwave), 530 nm (middlewave), MW, and 560 nm (longwave, LW) (12, 13). Color vision depends on a neural comparison of the quantum catches of the different cone classes. The photopigments are embedded in the multiply enfolded membranes of the cone outer segments, and each photopigment consists of a protein moiety bound to retinal, a derivative of vitamin A1. The proteins are members of the superfamily of

---

Fig. 1. Pelargonium zonale. A specimen from stock collected in South Africa that almost certainly corresponds to Dalton’s “Geranium zonale” (4).

Fig. 2. The preserved eye fragments of John Dalton, photographed in 1982.


*To whom correspondence should be addressed.
G protein–coupled receptor molecules, or heptahelicals (G proteins are guanosine triphosphate–binding proteins); they exhibit seven transmembrane helices, which are linked by intracellular and extracellular loops and form a palisade around the chromophore retinal (14).

Modern spectrophotometric measurements have confirmed that the dichromatic form of color blindness commonly arises from the lack of either the LW or the MW photopigment (Fig. 3) [the former condition is termed protanopia and the latter deuteranopia (12, 13)]. Molecular biology has shown that most of the inherited forms of color deficiency are associated with changes in the genes on the X chromosome that code for the protein moieties of the LW and MW visual pigments (15, 16). The LW and MW genes are both composed of six coding regions or exons, and such changes are thought to arise from illegitimate pairing between these highly homologous genes, followed by crossing-over (16).

Although there are 15 amino acid differences between MW and LW opsin genes (15), the greater part of the spectral shift in sensitivity between MW and LW visual pigments is the result of substitutions at sites 180, 277, and 285 (17–20), with five other sites having smaller effects (21). Site 180, however, is polymorphic (22, 23) in both the MW and LW opsin genes: The substitution of serine for alanine in an otherwise MW sequence increases the λmax by about 2 nm, and the substitution of alanine for serine in an otherwise LW sequence decreases the λmax by 4 to 7 nm (20, 21). We have used sites 277 and 285 (both coded by exon 5), together with gene-specific differences in exon 4, to determine whether Dalton possessed a MW or a LW opsin gene or both.

We took several small samples of tissue from Dalton’s preserved eye from regions tentatively identified as peripheral retina, and these samples were used for the amplification of opsin genes by the polymerase chain reaction (PCR) (24). Two different primer sets were used (25): One set is gene-specific (26) and amplifies either from MW or from LW exons, whereas the second set amplifies from both genes (18, 27, 28). The gene-specific set comprises a primer pair (R4+, R5–) that is specific to the LW exons 4 and 5 and another (G4+, G5–) that is specific to the MW exons 4 and 5 (Fig. 4). These primer pairs produce a fragment of about 1.7 kb. In other combinations (R4+, G5– and G4+, R5–), they generate fragments from any hybrid gene that contains a MW exon 4 and LW exon 5 or vice versa. Only primer pair R4+, R5– generated a fragment of the correct size from Dalton’s DNA (Fig. 5A), indicating that a LW but no MW opsin gene is present. Southern (DNA) blot analysis and probing of the PCR products with an exon 5 probe confirmed that only this set of primers amplified an opsin gene fragment (Fig. 5B). The specificity of these primers was confirmed with a number of test DNA samples (Fig. 5C). The DNA from a normal observer yielded fragments of the correct size only when the R4+, R5– and G4+, G5– combinations were used. DNA from a known anomalous trichromat produced, in addition to these fragments, a fragment with the R4+, G5– combination, indicating the presence of a hybrid gene. And lastly, only the R4+, R5– primer pair amplified a fragment from the DNA of a known deuteranope. The identity of the fragment amplified from Dalton’s DNA by the R4+, R5– primer pair was then confirmed by cloning and sequencing (29).

Partial sequencing of 10 separate clones (Fig. 6) show that the fragment was amplified from an opsin gene that coded for a LW visual pigment.

![Diagram](image-url)
We sought further evidence for the presence of only a LW opsin sequence in Dalton by using nonspecific primer pairs (18, 27, 28) that do not differentiate between MW and LW genes. These primers also provide information on codon 180 in exon 3, codons 230 and 233 in exon 4, and codons 277 and 285 in exon 5. Four pairs were used, three that amplify within exons 3, 4, and 5 respectively, and a fourth pair that amplifies from intron 4, across exon 5, and into intron 5 (Fig. 4). For each primer pair, the resulting amplified fragment was cloned and sequenced. In all cases, only a single sequence was amplified (Fig. 6): For exons 4 and 5, all of the clones were identical to the sequence of the LW opsin gene reported by Nathans and his colleagues (15), whereas all of the exon 3 clones coded for Ala180, the sequence being identical to the second most common variant of this highly polymorphic region of the LW opsin gene (23). A consistent result was therefore obtained from 27 different clones that contained fragments generated with a range of nonspecific primers that amplify from both intron and exon regions of the MW and LW opsin genes: Only a single LW opsin sequence is present in Dalton’s DNA.

Analyses of old DNA by PCR are bedeviled by contamination (30). For the following reasons, however, we do not believe that our result is artifactual. We deliberately took the samples from the interior of the eye, which is unlikely to have been handled. Samples were removed with a sterile scalpel blade and placed immediately into sterile 1.5-mL eppendorf tubes. At all times, sterile surgical gloves were worn. New pipettes and reagents were used in all subsequent procedures, and all experiments were set up in a room that had not been previously used for work on visual pigments. Each PCR used a different sample of eye tissue yet gave the same overall result. And there is a further unusual and salient argument against contamination in the present case: Our result depends on the absence, rather than the presence, of a gene—a gene whose LW fellow is present. Only contamination by the DNA from another deuteranope would have given the same result. However, as far as we are aware, no workers in our laboratory have any defect in color vision and certainly none are dichromats, and no samples of DNA from other dichromats were present in our laboratory until after the analysis of Dalton’s DNA was complete. The deuteranope that we used for test purposes was not known to us until after all of the PCR amplifications on Dalton’s tissue had been carried out.

Therefore, from our data we can confirm that Dalton was indeed a dichromat, but contrary to previous interpretations, he was a deuteranope with a single LW opsin gene that coded for a visual pigment with the shorter spectral sensitivity of the two common LW variants. Long tradition holds that Dalton was a protanope (11, 31). Can our result be reconciled with the historical evidence for his phenotype?

The classical sign of protanopia is the foreshortening of the red end of the spectrum because of the absence of the LW receptor. However, when the physicists Sir John Herschel and Sir David Brewster each questioned Dalton directly, they both reported that he did not see the spectrum as foreshortened at long wavelengths (8, 32). Similarly, Richard Taylor, the scientific
We showed a specimen of *P. zonale* (Fig. 1) to the deuteranope whose DNA was analyzed as a test for the gene-specific primer pairs (Fig. 5C). In northern daylight he described the color as similar to that of the sky, but slightly desaturated. In candlelight a few minutes later, he described the color as red.

**REFERENCES AND NOTES**

1. The lecture by Dalton was delivered on 31 October 1794. It was published in Mem. R. Soc. (1816) and reprinted in the Edinburgh J. Sci. 5, 186 (1831).

2. French, Spanish, and Russian are among the languages that have the term deutanopia. The term daltonian was used in print as early as 1857 by P. Prevost [Bibl. Univ. Sci. Belg. Arts Sci. Arts 35, 320 (1827)].


4. The "horseshoe cranefoot" or *Gerranium zonale* is illustrated by Thomas Martyn [Thirty-Eight Plates Intended to Illustrate Linnaeus's System (J. White, London, 1769)], and we can be fairly certain that the species to which Dalton refers is the one now known as *Gerranium zonale*, a native plant of South Africa imported to Britain during the 18th century.


8. C. Barrett, Diary and Letters of Madame D'Arblay (Macmillan, London, 1904). George III told the future Madame D'Arblay, "There are people who have no eye for difference of colour. The Duke of Marlborough actually cannot tell scarlet from green...I do not think that this defect runs in his family, for Lady Di Beauclair draws very finely."


10. "Die Umwirksamkeit der Molekülen...deren außerordentliche Beweglichkeit." This may be the earliest suggestion that a hereditary disorder was due to a molecular defect.


Vira!lly encoded human immunodeficiency virus (HIV) reverse transcriptase (RT) catalyzes the replication of single-stranded viral RNA to yield double-stranded DNA.

The mechanism of inhibition of HIV-1 reverse transcriptase by three nonnucleoside inhibitors is described. Nevirapine, O-TIBO, and CI-TIBO each bind to a hydrophobic pocket in the enzyme-DNA complex close to the active site catalytic residues. Pre-steady-state kinetic analysis was used to establish the mechanism of inhibition by these noncompetitive inhibitors. Analysis of the pre-steady-state burst of DNA polymerization indicated that inhibitors blocked the chemical reaction, but did not interfere with nucleotide binding or the nucleotide-induced conformational change. Rather, in the presence of saturating concentrations of the inhibitors, the nucleoside triphosphate bound tightly (Ki < 100 mM), but nonproductively. The data suggest that an inhibitor combining the functionalities of a nonnucleoside inhibitor and a nucleotide analog could bind very tightly and specifically to reverse transcriptase and could be effective in the treatment of AIDS.

before the viral genome is integrated into the DNA of the host. RT has been the target of several antiviral therapeutic agents (1) used in the treatment of AIDS. Although nucleoside analogs, such as AZT (3'-azido 3'-deoxythymidine) and dDC (2',3'-dideoxythymidine), have been the drugs of choice in attenuating the action of the virus, toxicity limits their use (2) and long-term inhibition of RT is limited by the high frequency of virus mutation whereby drug-resistant forms may accumulate (3). In searches for new inhibitors, random screening has resulted in the identification of several classes of nonnucleoside RT inhibitors. These include the tetrahydrobenzodiazepine (TIBO) derivatives and the dipyrindiazepine Nevirapine (4-6) (Fig. 1). Because the nonnucleoside derivatives inhibit RT at low concentrations and with high specificity, these inhibitors are promising candidates for the treatment of AIDS. However, as in the case of the nucleoside analogs (3, 7), TIBO- and Nevirapine-resistant forms of RT have already been identified (8, 9). Several studies undertaken in an effort to retard the onset of resistance have been focused on convergent combination therapy (2, 10).

A crystal structure of RT with Nevirapine bound shows the inhibitor in a hydrophobic pocket near the polymerization active site (11). Nevirapine makes contact with the side chain residues of Tyr (19) and Tyr in the p66 subunit only. The TIBO derivatives are structurally distinct classes of inhibitors that appear to bind to the Nevirapine site (12), a conclusion that is supported by the observed cross-resistance of these drugs in mutant RTs (8, 13). However, structural information cannot reveal which step or steps in the DNA polymerization cycle are affected by binding the inhibitor at a distant site.

Several steady-state studies have attempted to elucidate the inhibition mechanism of nonnucleoside inhibitors (6, 14, 15). In general, these studies suggest a noncompetitive inhibition pattern with respect to both primer-template and nucleotide substrates. However, the steady-state kinetic formalism does not provide an under-